Ath. Kehagias and M. Konstantinidou.
"L-Fuzzy Valued Inclusion Measure, L-Fuzzy Similarity and L-Fuzzy Distance".

This paper will appear in the journal:
Fuzzy Sets and Systems

L-Fuzzy Valued Inclusion Measure,

L-Fuzzy Similarity and L-Fuzzy Distance

Ath. Kehagias and M. Konstantinidou

Box 464, Division of Mathematics

Dept. of Mathematics, Physical and Computational Sciences,

Faculty of Engineering, Aristotle University of Thessaloniki

Thessaloniki, GR 54006, GREECE

email: kehagias@egnatia.ee.auth.gr

20 July 2001

Abstract

The starting point of this paper is the introduction of a new *measure of inclusion* of fuzzy set A in fuzzy set B. Previously used inclusion measures take values in the interval [0,1]; the inclusion measure proposed here takes values in a *Boolean lattice*. In other words, inclusion is viewed as an *L-fuzzy valued relation* between fuzzy sets. This relation is reflexive, antisymmetric and transitive, i.e. it is a *fuzzy order relation*; in addition it possesess a number of properties which various authors have postulated as axiomatically appropriate for an inclusion measure. We also define an L-fuzzy valued measure of *similarity* between fuzzy sets and and an L-fuzzy valued *distance* function between fuzzy sets; these possess properties analogous to the ones of real-valued similarity and distance functions.

Keywords: Fuzzy Relations, inclusion measure, subsethood, L-fuzzy sets, similarity, distance, transitivity.

1 Introduction

A *measure of inclusion* (also called a *subsethood measure*) is a *relation* between fuzzy sets A and B, which indicates the degree to which A is contained in (is a subset of) B. Many measures of inclusion have been proposed in the literature; these usually take values either in $\{0,1\}$ or in [0,1]. The first case corresponds to a *crisp* relation: A is either contained or not contained in B, while the second case corrresponds to a *fuzzy* relation: A is contained in B to a certain degree. A related concept is that of a *measure of similarity* between fuzzy sets; a *measure of similarity* is a relation which can be seen as a fuzzification of a crisp *equivalence relation*.

In this paper we introduce an inclusion measure, denoted by $I(A,B)$, i.e. $I(A,B)$ denotes the degree to which fuzzy set A is included in fuzzy set B. We also introduce a related similarity measure denoted by $S(A,B)$ and a *distance between fuzzy sets* $D(A,B)$. The main difference between our work and that of other authors is the following: traditionally, inclusion measures, similarities and distances take values in [0,1] or some other *totally ordered* subset of the *real* interval [0,∞); in this paper, on the other hand, the range of $I(.,.)$, $S(.,.)$ and $D(.,.)$ is a *Boolean lattice* $\mathbf{B}$. Since $\mathbf{B}$ is a *partially* (but not totally) ordered set, it follows that the proposed inclusion, similarity and distance are *L-fuzzy valued relations between fuzzy sets*. For the sake of brevity we will occasionally refer to quantities which take values in a totally ordered set as *"scalar"* and to quantities which take take values in a partially ordered set as *"vector"*. The rationale for this terminology will become obvious in the sequel.

The proposed inclusion measure $I(A,B)$ is reflexive, symmetric and *transitive* – hence is a *L-fuzzy order relation*. In addition, $I(A,B)$ possesess a number of attractive properties which various authors have postulated as axiomatically appropriate for an inclusion measure. Similar remarks hold for the proposed similarity $S(A,B)$ and distance $D(A,B)$.

Let us now briefly review related previous work. Zadeh, in the seminal paper [33], gives the first definition of fuzzy set inclusion; in Zadeh's formulation, inclusion is a *crisp* relation (i.e. a fuzzy set A is either included or not included in a fuzzy set B). Also, there is a popular concept of fuzzy set inclusion which is essentially a fuzzy analog of conditional probability (in this paper it appears as inclusion measure no.1 in Section 3.2). This appears in many works (see for instance [4]). An interesting variation of this measure (in this paper it appears as inclusion measure no.2 in Section 3.2) appears in [9, 30, 13, 14, 24].

Several authors take an *axiomatic* approach to the study of inclusion measure, i.e. they provide a list of properties ("axioms") which a "reasonable" inclusion measure must satisfy and then they examine whether a particular inclusion measure or a family of inclusion measures possesses these properties. A prime example of this approach is Sinha and Dougherty's [26]. They list nine properties that a "reasonable" inclusion measure should have and then proceed to introduce inclusion measures which have these properties.

A somewhat different point of view is taken by Bandler and Kohout in [2]. These authors obtain several inclusion measures from *fuzzy implication operators*. A related approach is that of [30, 31] where the transitivity of inclusion measure is also studied. The issue of transitivity is also studied in [17], where a $\wedge$–transitive inclusion measure is introduced.

Axiomatization and connection to implication operators are combined in a number of papers. Young combines the above approaches in [32] and also connects inclusion measure to fuzzy entropy. Young gives various examples of inclusion measures listed in the literature and examines which of the properties she proposes are satisfied by these inclusion measures. She also gives the connection to Bandler and Kohout's work on fuzzy implication operators. Finally, Young presents an application of inclusion measure to tuning fuzzy logic rules. In [6], Burillo et al. introduce a particular family of fuzzy implication operators (a form of generalized Lukasiewicz operators) and show that the inclusion measures obtained from this family (in Bandler and Kohout's manner), satisfy Sinha and Dougherty's axioms for inclusion measures. In [9] J. Fan et al. discuss the connections between inclusion measure, fuzzy entropy and fuzzy implication. They comment on Young's axioms and propose their own list of axioms. They also give some conditions to check if a function is an inclusion measure. Finally, they present an application of inclusion measure to clustering validity.

All the inclusion measures discussed up to this point take values in a totally ordered set. In [7], Bustince introduces an inclusion measure which takes values in the partially ordered set of interval valued fuzzy sets. Bustince also relates inclusion to implication operators.

A somewhat different view on inclusion measures is taken in [5]. The authors classify measures (of similarity, satisfiability, inclusion, dissimilarity etc.) using a very general scheme to generate such measures from elementary functions; the corresponding properties are also quite general. The authors also consider the *aggreggation* of several local measures of similarity into a global one.

In some of the above mentioned works, the connection between inclusion and similarity measures is discussed. Fuzzy similarity has received much more attention than fuzzy inclusion and the corresponding literature is very extensive; here we discuss a small number of papers, which have a point of view similar to ours.

Pappis and his collaborators have issued a series of papers [21, 22, 23] which take an axiomatic view to similarity measures. In [10] J. Fan gives axioms for entropy, distance and similarity. Liu in [18] proposes an alternative set of axioms. X. Wang et al. criticize Pappis' work in [29], present a modified definition of similarity and explore its connection to Bandler and Kohout's work. W. Wang in [28] adopts Liu's similarity axioms and introduces two new similarity measures. Ovchinnikov in [20] discusses L-fuzzy relations, i.e. relations taking values in lattices and mentions similarity relations in this context.

Returning to the topic of inclusion measures, we should mention [13, 14, 24] which take a somewhat different view towards inclusion. Namely, the authors introduce a fuzzy measure of the inclusion of a *crisp* set A in a *crisp* set B and develop an extensive methodology which uses their inclusion measure for clustering and classification applications.

As we have already pointed out, practically all of the work on fuzzy inclusion and similarity uses measures taking values in the *totally ordered* interval [0,1]. A notable exception is [7] (which uses interval valued similarity measures). A more general discussion of L-fuzzy valued relations can be found in, for example, [11] and [25].

2 Preliminaries

2.1 Fuzzy and Crisp Sets

All subsequent discussion makes use of two fundamental sets. First, we have a *universe of discourse* denoted by U; this can be quite general, i.e. no special structure or properties are assumed (in particular, U can be finite, countable or uncountable and we assume no order on the elements of U). Second, we have a *totally ordered* set denoted by L; in this paper we take $L = [0, 1]$ (but most of our results remain valid in case L is a finite set $\{a_1, a_2, ..., a_{2N+1}\}$ with $0 = a_1 < a_2 < ... < a_{N+1} = \frac{1}{2} < ... < a_{2N} < a_{2N+1} = 1$). Elements of both U and L will be denoted by lowercase letters. For elements

$x, y \in L$ we will use $x \leq y$ to denote that x is less than or equal to y; the symbols $<$, $\geq$, $>$ also have their usual significance. The minimum of x and y will be denoted by $x \wedge y$ and the maximum of x and y will be denoted by $x \vee y$. We also introduce negation of the elements of L (denoted by $'$) which is defined by: $a' \doteq 1 - a$; here "$-$" indicates the usual subtraction of real numbers. Hence $(L, \leq, \wedge, \vee, ')$ is a *de Morgan lattice*, i.e. a bounded lattice, *distributive* with respect to the $\wedge, \vee$ operations, *order-inverting* with respect to the $'$ operation and satisfying de Morgan's laws [19]. In addition, $(L, \leq)$ is a complete lattice and totally ordered.

We will be concerned with fuzzy (sub)sets of U; fuzzy sets are identified with their membership functions, which take values in L. In other words, a fuzzy set is a function $A : U \rightarrow L$. Fuzzy sets will be denoted by uppercase letters: $A, B, C, \ldots$, with two exceptions: the empty fuzzy set will be denoted by $\underline{0}$ (this is equivalent to $\emptyset$) and the universal fuzzy set will be denoted by $\underline{1}$ (this is equivalent to U). In case U is a denumerable set, then we can represent a fuzzy set A in vector notation $A = [A_1, A_2, \ldots]$. Hence the degree to which element u belongs in set A is denoted by A_u. The family of all fuzzy sets (the fuzzy *powerset* of U) will be denoted by $\mathbf{F}(U)$ or simply by $\mathbf{F}$.

Example 2.1 *Take $U = \{1, 2, 3, 4\}$. Then we have $\underline{1} = [1, 1, 1, 1]$ (i.e. $\underline{1}_1 = 1, \underline{1}_2 = 1, \underline{1}_3 = 1, \underline{1}_4 = 1$) and $\underline{0} = [0, 0, 0, 0]$. Other examples of fuzzy sets are $A = [0.2, 0.3, 0.0, 0.9]$ (i.e. $A_1 = 0.2, A_2 = 0.3, A_3 = 0.0, A_4 = 0.9$), $B = [0.3, 0.4, 0.2, 0.8]$, $C = [0.3, 0.5, 0.8, 1.0]$.*

Crisp (sub)sets of U are (obviously) special cases of fuzzy sets and are also identified with their membership (or characteristic) functions, which take values in $\{0, 1\}$. In other words, a fuzzy set is a function $\Theta : U \rightarrow \{0, 1\}$. To underscore the fact that a set is crisp we will usually denote it by an uppercase *Greek* letter such as: $\Theta, \Phi, \Xi, \ldots$. The family of all crisp sets (the crisp powerset of U) will be denoted by $\mathbf{B}(U)$ or simply by $\mathbf{B}$. Clearly $\mathbf{B} \subseteq \mathbf{F}$.

Example 2.2 *Take $U = \{1, 2, 3, 4\}$. Examples of crisp sets are $\Theta = [0, 1, 1, 1]$ and $\Phi = [1, 1, 0, 1]$.*

2.2 The $(\mathbf{F}, \leq)$ Lattice

We have already remarked that $(L, \leq, \wedge, \vee, ')$ is a totally ordered *de Morgan lattice*. We now define an order on elements of $\mathbf{F}$. This is obtained from the "elementwise" order and is denoted (without danger of confusion) by the same symbol $\leq$. I.e., for all $A, B \in \mathbf{F}$ we define

$$A \leq B \ \Leftrightarrow \ (\text{for all } u \in U \text{ we have } A_u \leq B_u).$$

It is easy to see that $\leq$ is an order on $\mathbf{F}$. More rigorously,

$$\mathbf{F} \doteq \underbrace{L \times L \times ... \times L}_{|U| \text{ times}}.$$

and $(\mathbf{F}, \leq)$ is the direct product lattice [1]

$$(\mathbf{F}, \leq) \doteq \underbrace{(L, \leq) \times (L, \leq) \times ... \times (L, \leq)}_{|U| \text{ times}}.$$

$(\mathbf{F}, \leq)$ is a *partially* ordered set, i.e. there exist elements $X, Y \in \mathbf{F}$ such that neither $X \leq Y$ nor $Y \leq X$; this is denoted by $X||Y$ (X and Y are *incomparable*). Furthermore, as a product of complete lattices, $(\mathbf{F}, \leq)$ is complete [8].

The sup and inf operations on elements of $\mathbf{F}$ (with respect to $\leq$) can be obtained from the elementwise min and max operations. I.e. for all $A, B \in \mathbf{F}$ the $\inf(A, B)$ exist; it is denoted (without danger of confusion) by $A \wedge B$ and defined for all $u \in U$ by

$$(A \wedge B)_u \doteq A_u \wedge B_u.$$

(Notice the notation above: $A \wedge B$ is a function; the value of the function at u is denoted by $(A \wedge B)_u$.) Similarly, for all $A, B \in \mathbf{F}$ the $\sup(A, B)$ exists; it is denoted (without danger of confusion) by $A \vee B$ and defined for all $u \in U$ by

$$(A \vee B)_u \doteq A_u \vee B_u.$$

Finally, complementation is defined on $\mathbf{F}$ in terms of elementwise complementation: for all $A \in \mathbf{F}$ and $u \in U$ we have $(A')_u = (A_u)'$.

Example 2.3 *We continue using the sets of the previous examples. It can be seen that for* $u = 1, 2, 3, 4$ *we have* $A_u \leq C_u$*; hence* $A \leq C$, $A \vee C = C$, $A \wedge C = A$. A *and* B *are incomparable:* $A||B$. *We have* $A \vee B = [0.3, 0.4, 0.2, 0.9]$ *and* $A \wedge B = [0.2, 0.3, 0.0, 0.8]$*; we write (for instance)* $(A \wedge B)_1 = 0.2$. *Similarly* $B \vee \Phi = [1, 1, 0.2, 1]$. *Finally,* $A' = [0.8, 0.7, 1.0, 0.1]$.

It is easy to prove that $(\mathbf{F}, \leq, \wedge, \vee, ')$ is a de Morgan lattice. It is also easy to prove that $(\mathbf{B}, \leq, \wedge, \vee, ')$ is a *Boolean lattice* (or *Boolean algebra*), a complete lattice and a sublattice of $(\mathbf{F}, \leq, \wedge, \vee, ')$.

2.3 Fuzzy Relations

In the sequel we will often use *fuzzy relations* on *subsets of* U (e.g. inclusion, similarity, distance). In this paper, the fuzzy relations of interest are functions of the form $R : \mathbf{F} \times \mathbf{F} \rightarrow \mathbf{B}$. Adapting well known definitions [15] to this context we have the following.

Definition 2.4 *A fuzzy relation* R *is called* reflexive *iff* $\forall A \in \mathbf{F}$ *we have* $R(A, A) = \underline{1}$.

Definition 2.5 *A fuzzy relation* R *is called* symmetric *iff* $\forall A, B \in \mathbf{F}$ *we have* $R(A, B) = R(B, A)$.

Definition 2.6 *A fuzzy relation* R *is called* antisymmetric *iff* $\forall A, B \in \mathbf{F}$ *we have*

$$R(A, B) = R(B, A) > \underline{0} \Rightarrow A = B.$$

Definition 2.7 *A fuzzy relation* R *is called* $\wedge$-transitive *iff* $\forall A, B \in \mathbf{F}$ *we have*

$$R(A, B) \geq \sup_{C \in \mathbf{F}} [R(A, C) \wedge R(C, B)].$$

Definition 2.8 *A fuzzy relation* R *is called* $\vee$-transitive *iff* $\forall A, B \in \mathbf{F}$ *we have*

$$R(A, B) \leq \inf_{C \in \mathbf{F}} [R(A, C) \vee R(C, B)].$$

Remark. In the above definitions the indicated sup and inf exist always because $(\mathbf{B}, \leq, \wedge, \vee, ')$ is a complete lattice.

Definition 2.9 *A fuzzy relation R is called a* fuzzy similarity relation *(or* fuzzy equivalence relation*) if it is reflexive, symmetric and $\wedge$-transitive.*

Definition 2.10 *A fuzzy relation R is called a* fuzzy order relation *if it is reflexive, antisymmetric and $\wedge$-transitive.*

Note that $R(A, B)$ is a function with *domain* $\mathbf{F} \times \mathbf{F}$ and *range* $\mathbf{B}$.To denote the value of $R(A, B)$ at point $u \in U$, we use the notation $R_u(A, B)$.[1]

3 The L-Fuzzy Inclusion Measure

3.1 Definitions and Properties

Definition 3.1 *For all $A, B \in \mathbf{F}$, the* measure of inclusion *of A in B is a fuzzy relation, denoted by $I(A, B)$. The value of $I(A, B)$ is defined for each $u \in U$ by*

$$I_u(A, B) \doteq \begin{cases} 1 & \text{iff } A_u \leq B_u \\ 0 & \text{else.} \end{cases}$$

Example 3.2 *Continuing with the sets of the previous examples, we have: $I(A, C) = [1, 1, 1, 1]$, $I(A, B) = [1, 1, 1, 0]$, $I_1(A, B) = 1$, $I_4(A, B) = 0$. Also, $I(A, \Phi) = [1, 1, 1, 1]$, $I(\Theta, A) = [1, 0, 0, 0]$, and $I(B, A) = [0, 0, 0, 1]$. Note that $I(A, B) \leq I(A, C)$, but $I(\Theta, A) || I(B, A)$.*

The motivation behind the above definition is as follows. Inclusion of set A into set B is defined in terms of the elementwise values of A and B. For example, if for every $u \in U$ it is true that u is contained in B more than it is contained in A, then we conclude that A is included in B in the maximum degree, namely $\underline{1}$. The difference of Definition 3.1 from alternative definitions of inclusion measure is the following: traditionally the "elementwise" inclusion of two sets is aggregated into a single real number, i.e. the range of a traditional inclusion measure is usually totally ordered; whereas, in this work the elementwise inclusions are collected into $I(A, B)$, a quantity which takes values in the

[1] It is slightly unusual to place the argument of a function as a subscript, but this notation will prove advantageous in the sequel.

partially ordered set $\mathbf{B}$ (i.e. a L-fuzzy quantity). Since the values of $I(.,.)$ are partially but not totally ordered, there may exist sets X, Y, Z, W such that: X is included in Z more than it is included in Y (i.e. $I(X,Y) \leq I(X,Z)$) but the inclusion of X in Y *cannot be compared* to the inclusion of X in W (i.e. $I(X,Y)||I(X,W)$, viz. the last part of Example 3.2 above).

As mentioned in the introduction, we will occasionally refer to quantities which take values in a totally ordered set as *"scalar"* and to quantities which take take values in a partially ordered set as *"vector"*. The rationale is rather obvious: in case U has finite cardinality ($|U| = N < \infty$), then $I(A,B)$ is indeed a Boolean vector of size N. In case U has infinite (countable or uncountable) cardinality, $I(A,B)$ is a function, which can still be considered as a vector in an infinite-dimensional vector space. The term "vector" is not intended to invoke connotations of vectorial operations, e.g. addition of inclusion measures; on the other hand, the partial order $\leq$ and the $\vee$ and $\wedge$ operations defined in Section 2.2 are the ones "naturally" introduced for vectors. The term "scalar" is used in contradistinction to "vector".

The following proposition describes what we consider the most attractive property of $I(.,.)$.

Proposition 3.3 *$I(.,.)$ is a fuzzy order on $\mathbf{F}$, i.e. it is a reflexive, antisymmetric and $\wedge$-*transitive *fuzzy relation*

Proof. Take any $A, B, C \in \mathbf{F}$. We then have the following.

1. *Reflexivity.* We have: $(\forall u \in U : A_u \leq A_u) \Rightarrow (\forall u \in U : I_u(A,A) = 1) \Rightarrow I(A,A) = \underline{1}$.

2. *Antisymmetry.* $I(A,B) = I(B,A) \Rightarrow (\forall u \in U : I_u(A,B) = I_u(B,A))$. Take any $u \in U$ and consider two cases. (a) If $I_u(A,B) = I_u(B,A) = 1$, then $A_u \leq B_u$ and $B_u \leq A_u$, hence $A_u = B_u$. (b) If $I_u(A,B) = I_u(B,A) = 0$, then $A_u > B_u$ and $B_u > A_u$ which leads to a contradiction. Hence we have that: $(\forall u \in U : I_u(A,B) = I_u(B,A)) \Rightarrow I(A,B) = I(B,A) = \underline{1} \Rightarrow A = B$.

3. $\wedge$-*transitivity.* We take any $u \in U$ and consider two cases. (a) If $I_u(A,C) \wedge I_u(C,B) = 0$, then clearly $I_u(A,B) \geq I_u(A,C) \wedge I_u(C,B)$. (b) If $I_u(A,C) \wedge I_u(C,B) = 1$, then $A_u \leq C_u$ and $C_u \leq B_u$, which implies $A_u \leq B_u$ and so $I_u(A,B) = 1 = I_u(A,C) \wedge I_u(C,B)$. Hence we have that: $(\forall u \in U : I_u(A,B) \geq I_u(A,C) \wedge I_u(C,B)) \Rightarrow I(A,B) \geq I(A,C) \wedge I(C,B) \Rightarrow I(A,B) \geq \sup_{C \in \mathbf{F}} I(A,C) \wedge I(C,B)$■

We now proceed to show that $I(.,.)$ has a number of additional properties, which several authors have axiomatically postulated as appropriate and/or desirable for measures of inclusion.

Proposition 3.4 *For all $A, B, C, D \in \mathbf{F}$ we have:*

I1 $I(A,B) = \underline{1} \Leftrightarrow A \leq B$; $I(A,B) = \underline{0} \Leftrightarrow A > B$.

I2 $I(A,A') = \underline{0} \Leftrightarrow (\forall u \in U$ we have $A_u > 1/2)$.

I3 $I(A,B) = I(B',A')$.

I4 $I(A,B) \vee I(B,A) = \underline{1}$.

I5 $B \leq C \Rightarrow I(A,B) \leq I(A,C)$.

I6 $B \leq C \Rightarrow I(C,A) \leq I(B,A)$.

I7 $I(A,B) \wedge I(C,D) \leq I(A \wedge C, B \wedge D) \wedge I(A \vee C, B \vee D) \leq$

$$I(A \wedge C, B \wedge D) \vee I(A \vee C, B \vee D) \leq I(A,B) \vee I(C,D).$$

I8 $I(A \vee B, C) = I(A,C) \wedge I(B,C)$.

I9 $I(A \wedge B, C) = I(A,C) \vee I(B,C)$.

I10 $I(A, B \vee C) = I(A,B) \vee I(A,C)$.

I11 $I(A, B \wedge C) = I(A,B) \wedge I(A,C)$.

I12 $I(A,B) \leq I(A \wedge C, B \wedge C) \wedge I(A \vee C, B \vee C)$.

Proof. Take any $A, B, C, D \in \mathbf{F}$. In proving properties **I1–I11**, we will frequently make use of the fact that: for any $u \in U$ the range of A_u, B_u etc. is [0,1], which is a *totally ordered* set. This implies, for instance, that either $A_u \leq B_u$ or $A_u > B_u$ (which is generally not true in a partially ordered set); this and similar facts will be of key importance in several of the arguments presented in the sequel.

I1 Assume $I(A,B) = \underline{1}$. Then we have: $(\forall u \in U : I_u(A,B) = 1) \Leftrightarrow (\forall u \in U : A_u \leq B_u) \Leftrightarrow A \leq B$.

In a similar manner we can prove $I(A,B) = \underline{0} \Leftrightarrow A > B$.

I2 Assume $I(A, A') = \underline{0}$. Then for all $u \in U$ we have $I_u(A, A') = 0 \Leftrightarrow A_u > A'_u \Leftrightarrow A_u > 1 - A_u \Leftrightarrow A_u > 1/2$.

I3 We take any $u \in U$ and consider two cases. (a) For all $u \in U$ such that $I_u(A, B) = 1$ we have $A_u \leq B_u \Rightarrow B'_u \leq A'_u \Rightarrow I_u(B', A') = 1$. (b) For all $u \in U$ such that $I_u(A, B) = 0$ we have $A_u > B_u \Rightarrow B'_u > A'_u \Rightarrow I_u(B', A') = 0$. So, we have $(\forall u \in U : I_u(A, B) = I_u(B', A')) \Rightarrow I(A, B) = I(B', A')$.

I4 We take any $u \in U$. Then either $A_u \leq B_u$ or $A_u > B_u$; so either $I_u(A, B) = 1$ or $I_u(B, A) = 1$. Hence we have: $(\forall u \in U : I_u(A, B) \vee I_u(B, A) = 1) \Rightarrow I(A, B) \vee I(B, A) = \underline{1}$.

I5 Assume $B \leq C$. Then for all $u \in U$ we have $B_u \leq C_u$. Now we take any $u \in U$ and consider two cases. (a) If $A_u > B_u$, then $I_u(A, B) = 0 \leq I_u(A, C)$. (b) If $A_u \leq B_u$, then $A_u \leq C_u \Rightarrow I_u(A, B) = 1 = I_u(A, C)$. Hence we have: $(\forall u \in U : I_u(A, B) \leq I_u(A, C)) \Rightarrow I(A, B) \leq I(A, C)$.

I6 Assume $B \leq C$. For all $u \in U$ we have $B_u \leq C_u$. Now we take any $u \in U$ and consider two cases. (a) $C_u > A_u$, then $I_u(C, A) = 0 \leq I_u(B, A)$; (b) $C_u \leq A_u$, then $B_u \leq A_u \Rightarrow I_u(C, A) = 1 = I_u(B, A)$. Hence we have: $(\forall u \in U : I_u(C, A) \leq I_u(B, A)) \Rightarrow I(C, A) \leq I(B, A)$.

I7 Take any $u \in U$. If $I_u(A, B) = 0$ or $I_u(C, D) = 0$ then $I_u(A, B) \wedge I_u(C, D) = 0 \leq I_u(A \wedge C, B \wedge D) \wedge I_u(A \vee C, B \vee D)$. If, on the other hand, $I_u(A, B) = I_u(C, D) = 1$, then

$$\left.\begin{array}{c} A_u \leq B_u \\ C_u \leq D_u \end{array}\right\} \Rightarrow \left\{\begin{array}{c} A_u \wedge C_u \leq B_u \wedge D_u \\ A_u \vee C_u \leq B_u \vee D_u \end{array}\right\} \Rightarrow \left\{\begin{array}{c} I_u(A \wedge C, B \wedge D) = 1 \\ I_u(A \vee C, B \vee D) = 1. \end{array}\right.$$

Hence we have: $(\forall u \in U : I_u(A, B) \wedge I_u(C, D) \leq I_u(A \wedge C, B \wedge D) \wedge I_u(A \vee C, B \vee D)) \Rightarrow$

$$I(A, B) \wedge I(C, D) \leq I(A \wedge C, B \wedge D) \wedge I(A \vee C, B \vee D). \tag{1}$$

Clearly we also have

$$I(A \wedge C, B \wedge D) \wedge I(A \vee C, B \vee D) \leq I(A \wedge C, B \wedge D) \vee I(A \vee C, B \vee D). \tag{2}$$

Finally, take any $u \in U$. If $I_u(A,B) \vee I_u(C,D) = 1$, then obviously $I_u(A \wedge C, B \wedge D) \vee I_u(A \vee C, B \vee D) \leq I_u(A,B) \vee I_u(C,D)$. If, on the other hand, $I_u(A,B) \vee I_u(C,D) = 0$, then $I_u(A,B) = I_u(C,D) = 0$, and so

$$\left.\begin{array}{c} A_u > B_u \\ C_u > D_u \end{array}\right\} \Rightarrow \left\{\begin{array}{c} A_u \wedge C_u > B_u \wedge D_u \\ A_u \vee C_u > B_u \vee D_u \end{array}\right\} \Rightarrow \left\{\begin{array}{c} I_u(A \wedge C, B \wedge D) = 0 \\ I_u(A \vee C, B \vee D) = 0. \end{array}\right.$$

Hence we have: $(\forall u \in U : I_u(A \wedge C, B \wedge D) \vee I_u(A \vee C, B \vee D) \leq I_u(A,B) \vee I_u(C,D)) \Rightarrow$

$$I(A \wedge C, B \wedge D) \vee I(A \vee C, B \vee D) \leq I(A,B) \vee I(C,D). \tag{3}$$

Now **I7** is immediately obtained from (1-3).

I8 In (1) substitute B with C, C with B and D with C, to obtain

$$I(A,C) \wedge I(B,C) \leq I(A \wedge B, C) \wedge I(A \vee B, C) \leq I(A \vee B, C). \tag{4}$$

On the other hand, from **I6** we have

$$\left.\begin{array}{c} A \leq A \vee B \Rightarrow I(A \vee B, C) \leq I(A,C) \\ B \leq A \vee B \Rightarrow I(A \vee B, C) \leq I(B,C) \end{array}\right\} \Rightarrow I(A \vee B, C) \leq I(A,C) \wedge I(B,C). \tag{5}$$

Now (4,5) yield the desired result.

I9 In (3) substitute B with C, C with B and D with C, to obtain

$$I(A \wedge B, C) \leq I(A \wedge B, C) \vee I(A \vee B, C) \leq I(A,C) \vee I(B,C). \tag{6}$$

On the other hand, from **I5** we have

$$\left.\begin{array}{c} A \wedge B \leq A \Rightarrow I(A,C) \leq I(A \wedge B, C) \\ A \wedge B \leq B \Rightarrow I(B,C) \leq I(A \wedge B, C) \end{array}\right\} \Rightarrow I(A,C) \vee I(B,C) \leq I(A \wedge B, C). \tag{7}$$

Now (6,7) yield the desired result.

I10 In (3) substitute C with A and D with C, to obtain

$$I(A, B \vee C) \leq I(A, B \wedge C) \vee I(A, B \vee C) \leq I(A, B) \vee I(A, C). \tag{8}$$

On the other hand, from **I5** we have

$$\left.\begin{array}{l} B \leq B \vee C \Rightarrow I(A, B) \leq I(A, B \vee C) \\ C \leq B \vee C \Rightarrow I(A, C) \leq I(A, B \vee C) \end{array}\right\} \Rightarrow I(A, B) \vee I(A, C) \leq I(A, B \vee C). \tag{9}$$

Now (8,9) yield the desired result.

I11 In (1) substitute C with A and D with C, to obtain

$$I(A, B) \wedge I(A, C) \leq I(A, B \wedge C) \wedge I(A, B \vee C) \leq I(A, B \wedge C). \tag{10}$$

On the other hand, from **I6** we have

$$\left.\begin{array}{l} B \wedge C \leq B \Rightarrow I(A, B \wedge C) \leq I(A, B) \\ B \wedge C \leq C \Rightarrow I(A, B \wedge C) \leq I(A, C) \end{array}\right\} \Rightarrow I(A, B \wedge C) \leq I(A, B) \wedge I(A, C). \tag{11}$$

Now(10,11) yield the desired result.

I12 In (1) substitute D with C to get

$$I(A, B) \wedge I(C, C) \leq I(A \wedge C, B \wedge C) \wedge I(A \vee C, B \vee C). \tag{12}$$

But $I(C, C) = \underline{1}$ and so $I(A, B) \wedge I(C, C) = I(A, B)$. Now (12) yields **I12** immediately.■

Remark. Properties **I5, I6** express the monotonicity of $I(.,.)$ in the second and first argument, respectively. Properties **I8–I11** are related to the combination of this monotonicity and $\wedge$–transitivity; it is worth remarking that for inclusion measures which do not enjoy $\wedge$–transitivity, the corresponding properties are weaker, with $\leq$ or $\geq$ in place of $=$ (compare with the properties listed in Section 3.2).

Proposition 3.5 *For all $A, B, C, D \in \mathbf{F}$ and all $\Theta \in \mathbf{B}$ we have:*

$$I(A,B) \geq \Theta \Rightarrow I(A \wedge C, B \wedge C) \wedge I(A \vee C, B \vee C) \geq \Theta.$$

Proof. This follows immediately from **I12**. ■

3.2 Comparison to Previously Proposed Inclusion Measures

In this section we briefly compare our inclusion measure to the ones introduced by other authors. We will list some of the axioms/properties proposed by several authors as appropriate for a reasonable inclusion measure and will compare these to the properties of $I(.,.)$ introduced in this paper. We consider three axiomatizations of inclusion measure properties: Sinha and Dougherty's ([26], also reproduced in [6]), Fan's [9] and Young's [32]. We present these axioms in Table 1. However, before proceeding some remarks are in order.

First of all, let us note that we have adapted the notation of the authors mentioned above so as to parallel our notation. In particular, we have used the symbol $i(.,.)$ to denote a scalar inclusion measure (in contradistinction to our vector $I(.,.)$); also, since the axioms/ properties appearing in Table 1 refer to inclusion measures with a totally ordered range, the maximum element is 1 rather than $\underline{1}$ and the minimum element is 0 rather than $\underline{0}$.

Second, we must remark that some of the authors previously mentioned present several alternative lists of axioms/ properties; hence Table 1 is meant to give a representative (but not exclusive) view of what are considered "reasonable" properties for an inclusion measure.

Finally, let us note that the term "axiom" (even though used by some authors) is not entirely appropriate, because some of the properties listed in Table 1 are not independent of each other. For instance, **i4a** implies **i8** and **i5a** implies **i7**. Also, it can be seen that **i4a** is a stronger form of **i4b** and **i5a** is a stronger form of **i5b.** Similarly, we have altered the order of presentation of the axioms/ properties so as to agree with the one we have used.

The first column in Table 1 lists an id. number and the the final column lists the corresponding property of our inclusion measure, as appearing in Proposition 3.4. The third column list the actual axiom / property and the remaining columns indicate the papers which use the particular property.

No.	Property	[6, 26]	[9]	[32]	Corresponding in this paper
i1	$A \leq B \Leftrightarrow i(A, B) = 1$	$\times$	$\times$	$\times$	**I1**
i2a	$i(A, A') = 0 \Leftrightarrow A$ is crisp set			$\times$	**I2**
i2b	$(\forall u \in U : \frac{1}{2} \leq A_u) \Rightarrow (i(A, A') = 0 \Leftrightarrow A = U)$		$\times$		**I2**
i3	$i(A, B) = i(B', A')$	$\times$	$\times$		**I3**
i4a	$B \leq C \Rightarrow i(A, B) \leq i(A, C)$	$\times$		$\times$	**I5**
i4b	$B \leq C \leq A \Rightarrow i(A, B) \leq i(A, C)$		$\times$		**I5**
i5a	$B \leq C \Rightarrow i(C, A) \leq i(B, A)$	$\times$		$\times$	**I6**
i5b	$A \leq B \leq C \Rightarrow i(C, A) \leq i(B, A)$		$\times$		**I6**
i6	$i(A \vee B, C) = i(A, C) \wedge i(B, C)$	$\times$			**I8**
i7	$i(A \wedge B, C) \geq i(A, C) \vee i(B, C)$	$\times$			**I9**
i8	$i(A, B \vee C) \geq i(A, B) \vee i(A, C)$	$\times$			**I10**
i9	$i(A, B \wedge C) = i(A, B) \wedge i(A, C)$	$\times$			**I11**
i10	$i(A, B) = 0 \Leftrightarrow \exists u \in U : A_u = 1, B_u = 0$	$\times$			**I1**

Table 1: Inclusion Measure Axioms

It can be seen from the above table that for every property appearing in Table 1, our $I(.,.)$ enjoys an analogous (and sometimes stronger) property. While Table 1 is not an exhaustive list of all properties proposed as reasonable for an inclusion measure, it offers a pretty complete coverage of such properties; the ones omitted are not, in our judgement, crucial[2]. In this connection, Sinha and Dougherty's remarks (pp.19-20 of [26]) regarding the choice of appropriate "axioms" are particularly relevant.

In conclusion, our $I(.,.)$ appears to be a rather reasonable inclusion measure in the sense that it has a larg enumber of reasonable properties.

In addition $I(.,.)$ enjoys $\wedge$–transitivity. This is usually *not* the case with the scalar inclusion measures appearing in the literature. This can be seen by considering the following list of specific examples of such inclusion measures. In the following list we use the popular notation $|A| = \sum_{u \in U} A_u$.

[2]For instance, we have omitted Sinha and Dougherty's property about invariance of inclusion measure with respect to the shift operation; but a similar property can be easily proved with respect to relabelings of set elements.

1. $i(A, B) = |A \wedge B| / |A|$.

2. $i(A, B) = |B| / |A \vee B|$.

3. $i(A, B) = |A' \wedge B'| / |B'|$.

4. $i(A, B) = |A'| / |A' \vee B'|$.

5. $i(A, B) = |A' \vee B| / |A' \vee A \vee B \vee B'|$.

6. $i(A, B) = |A' \wedge A \wedge B \wedge B'| / |A \wedge B'|$.

7. $i(A, B) = (|A'| \vee |B|) / |A' \vee A \vee B \vee B'|$.

8. $i(A, B) = |A' \wedge A \wedge B \wedge B'| / (|A| \wedge |B'|)$.

9. $i(A, B) = \frac{\sum_{u \in U} 1 \wedge (1 - A_u + B_u)}{|U|}$.

10. $i(A, B) = \frac{\sum_{u \in U} (1 - A_u) \vee B_u}{|U|}$.

11. $i(A, B) = \frac{\sum_{u \in U} (1 - A_u + A_u B_u)}{|U|}$.

12. $i(A, B) = \sup\{\alpha : \forall u \in U \text{ we have } A_u \wedge \alpha \leq B_u \wedge \alpha\}$.

Inclusion measure no.1 above is Kosko's inclusion measure [16]; no. 12 is introduced in [17] and does not appear very widely in the literature; for references to the remaining inclusion measures see [9] and [32].

The only inclusion measure in the above list which is $\wedge$-transitive is no.12, Kundu's inclusion measure. In addition, Willmott introduces some transitive fuzzy inclusion measures in [30], which however involve a rather drastically modified definition of transitivity. In conclusion, it appears that $\wedge$–transitivity is a requirement which cannot be easily satisfied by an inclusion measure which takes values in a totally ordered range ("scalar" inclusion measure). On the other hand it is satisfied (in addition with a large number of other properties) by our L-fuzzy inclusion measure.

4 Similarity and Distance Defined in Terms of L-Fuzzy Inclusion

We now proceed to define L-fuzzy similarity and L-fuzzy distance in terms of L-fuzzy inclusion.

4.1 L-Fuzzy Similarity

Definition 4.1 *For all $A, B \in \mathbf{F}$, the* measure of similarity *between A and B is denoted by $S(A, B)$ and is defined by*

$$S(A, B) \doteq I(A, B) \wedge I(B, A).$$

The domain of $S(.,.)$ is the partially ordered set $\mathbf{B}$, and this results to possibly incomparable similarity measures.

Example 4.2 *Continuing with the sets of the previous examples, we have:* $S(\Phi, \Theta) = I(\Phi, \Theta) \wedge I(\Theta, \Phi) = [0, 1, 1, 1] \wedge [1, 1, 0, 1] = [0, 1, 0, 1]$*; and* $S(\Phi, A) = I(\Phi, A) \wedge I(A, \Phi) = [0, 0, 1, 0] \wedge [1, 1, 1, 1] = [0, 0, 1, 0]$.

The rationale of the above definition is as follows. If for two sets A and B we have that A is included in B to the maximum degree and vice versa, then the sets are identical and have maximum similarity. In fact, the basic property is given by the following proposition.

Proposition 4.3 *For all $A, B \in \mathbf{F}$ and for all $u \in U$ we have: $S_u(A, B) = 1 \Leftrightarrow A_u = B_u$.*

Proof. Take any $A, B \in \mathbf{F}$ and any $u \in U$. $S_u(A, B) = 1 \Leftrightarrow (I_u(A, B) = 1$ and $I_u(B, A) = 1) \Leftrightarrow (A_u \leq B_u$ and $B_u \leq A_u) \Leftrightarrow A_u = B_u$.■

Further properties are presented in Propositions 4.4, 4.5, 4.6.

Proposition 4.4 *$S(.,.)$ is a fuzzy similarity relation on $\mathbf{F}$, i.e. it is reflexive, symmetric and $\wedge$-transitive.*

Proof. Take any $A, B, C \in \mathbf{F}$. Then we have the following.

1. *Reflexivity.* From Proposition 3.3 we have that $I(A, A) = \underline{1}$. Since $S(A, A) = I(A, A) \wedge I(A, A)$ it follows $S(A, A) = \underline{1}$.

2. *Symmetry.* $S(A, B) = I(A, B) \wedge I(B, A) = I(B, A) \wedge I(A, B) = S(B, A)$.

3. *$\wedge$-transitivity.* Set $\Theta \doteq I(A, B)$, $\Psi \doteq I(B, A)$, $\Phi \doteq I(A, C)$, $\Gamma \doteq I(C, A)$, $\Omega \doteq I(C, B)$, $\Delta \doteq I(B, C)$. From Proposition 3.3 we have $\Theta \geq \Phi \wedge \Omega$, $\Psi \geq \Delta \wedge \Gamma$. Hence $\Theta \wedge \Psi \geq (\Phi \wedge \Omega) \wedge (\Gamma \wedge \Delta) \Rightarrow \Theta \wedge \Psi \geq (\Phi \wedge \Gamma) \wedge (\Omega \wedge \Delta) \Rightarrow S(A, B) \geq S(A, C) \wedge S(C, B) \Rightarrow S(A, B) \geq \sup_{C \in \mathbf{F}}(A, C) \wedge S(C, B)$.■

Proposition 4.5 *For all* $A, B, C, D \in \mathbf{F}$ *we have:*

S1 $S(A,B) = \underline{1} \Leftrightarrow A = B$; $S(A,B) = \underline{0} \Rightarrow A \neq B$.

S2a $S(A,A') = \underline{0} \Leftrightarrow (\forall u \in U$ *we have* $A_u \neq 1/2)$.

S2b $S(A,A') = \underline{1} \Leftrightarrow (\forall u \in U$ *we have* $A_u = 1/2)$.

S3 $S(A,B) = S(A',B')$.

S4 $A \leq B \leq C \Rightarrow \begin{cases} S(A,B) \geq S(A,C) \\ S(B,C) \geq S(A,C) \\ S(A,C) = S(A,B) \wedge S(B,C). \end{cases}$

S5 $S(A,B) \wedge S(C,D) \leq \begin{cases} S(A \vee C, B \vee D) \\ S(A \wedge C, B \wedge D). \end{cases}$

S6 $S(A,B) \wedge S(A,C) \leq \begin{cases} S(A, B \vee C) \\ S(A, B \wedge C). \end{cases}$

S7 $S(A,C) \vee S(B,C) \geq \begin{cases} S(A \vee B, C) \\ S(A \wedge B, C). \end{cases}$

S8 $S(A,B) \leq \begin{cases} S(A \vee C, B \vee C) \\ S(A \wedge C, B \wedge C). \end{cases}$

S9a $S(A, A \vee B) = S(B, A \wedge B)$.

S9b $S(A, A \wedge B) = S(B, A \vee B)$.

S10 *All of the following quantities are equal:*

$$\begin{array}{llll} (i) & S(A,B) & (iv) & S(A \wedge B, A \vee B) \\ (ii) & S(A \wedge B, A) \wedge S(A, A \vee B) & (v) & S(A, A \vee B) \wedge S(B, A \vee B) \\ (iii) & S(A \wedge B, B) \wedge S(B, A \vee B) & (vi) & S(A, A \wedge B) \wedge S(B, A \wedge B) \end{array} .$$

Proof. Take any $A, B, C, D \in \mathbf{F}$. Then we have the following.

S1 $S(A,B) = \underline{1} \Leftrightarrow\ I(A,B) = I(B,A) = \underline{1} \Leftrightarrow (I(A,B) = \underline{1}$ and $I(B,A) = \underline{1}) \ \Leftrightarrow\ (A \leq B$ and $B \leq A) \Leftrightarrow$ $A = B$. Similarly we can prove $S(A,B) = \underline{0} \Rightarrow A \neq B$.

S2a $S(A,A') = \underline{0} \Rightarrow\ (\forall\ u \in U:\ I_u(A,A') \wedge I_u(A',A) = 0) \Rightarrow\ \ (\forall\ u \in U:\ A_u > 1 - A_u$ or $1 - A_u > A_u) \Rightarrow$ $(\forall\ u \in U\ :\ A_u \neq 1/2$. Conversely, assume that for all $u \in U$ we have $A_u \neq 1/2$. It is easy to see that if, for example, $A_u < 1/2$ then $I_u(A',A) = 0$; similarly if $A_u > 1/2$ then $I_u(A,A') = 0$. Hence we have: $(\forall u \in U : I_u(A',A) \wedge I(A,A^{'}) = 0) \Rightarrow (\forall u \in U : S_u(A,A^{'}) = 0) \Rightarrow S(A,A') = \underline{0}$.

S2b $S(A,A') = \underline{1} \Leftrightarrow I(A,A') \wedge I(A',A) = \underline{1} \Leftrightarrow (\forall\ u \in U\ :\ I_u(A,A') = I_u(A',A) = 1) \Leftrightarrow (\forall\ u \in U\ :$ $A_u = A'_u = 1/2)$.

S3 $S(A,B) = I(A,B) \wedge I(B,A) =\ \ I(B',A') \wedge I(A',B') = S(B',A') = S(A',B')$.

S4 Take any $u \in U$ and consider two cases. (a) $S_u(A,C) = 0 \leq S_u(A,B)$. (b) $S_u(A,C) = 1$. Then $A_u = C_u$ and this, in conjunction to $A_u \leq B_u \leq C_u$, implies $A_u = B_u \Rightarrow S_u(A,B) = 1$ $= S_u(A,C)$. Hence we have: $(\forall u \in U : S_u(A,B) \geq S_u(A,C)) \Rightarrow S(A,B) \geq S(A,C)$. It can be proved similarly that $S(B,C) \geq S(A,C)$. Finally, from $S(A,B) \geq S(A,C)$ and $S(B,C) \geq$ $S(A,C)$ follows that $S(A,B) \wedge S(B,C) \geq S(A,C)$; but from transitivity we also have $S(A,B) \wedge$ $S(B,C) \leq S(A,C)$ and so $S(A,B) \wedge S(B,C) = S(A,C)$.

S5 From Proposition 3.4, **I7** we have

$$\left.\begin{array}{l} I(A,B) \wedge I(C,D) \leq I(A \wedge C, B \wedge D) \\ I(B,A) \wedge I(D,C) \leq I(B \wedge D, A \wedge C) \end{array}\right\} \Rightarrow$$

$$I(A,B) \wedge I(C,D) \wedge I(B,A) \wedge I(D,C) \leq I(A \wedge C, B \wedge D) \wedge I(B \wedge D, A \wedge C) \Rightarrow$$

$$S(A,B) \wedge S(C,D) \leq S(A \wedge C, B \wedge D).$$

Similarly we can prove $S(A,B) \wedge S(C,D) \leq S(A \vee C, B \vee D)$.

S6 This is obtained from **S5** by substituting C with A and D with C.

S7 From Proposition 3.4, **I8** and **I10** imply

$$\left.\begin{array}{l} I(A \vee B, C) = I(A, C) \wedge I(B, C) \\ I(C, A \vee B) = I(C, A) \vee I(C, B) \end{array}\right\} \Rightarrow$$

$$S(A\vee B, C) = I(A\vee B, C)\wedge I(C, A\vee B) = (I(A, C)\wedge I(B, C))\wedge(I(C, A)\vee I(C, B)) =$$

$$(I(A, C)\wedge I(B, C)\wedge I(C, A))\vee(I(A, C)\wedge I(B, C)\wedge I(C, B)) \leq$$

$$(I(A, C)\wedge(I(C, A)\vee(I(B, C)\wedge I(C, B)) = S(A, C)\vee S(B, C).$$

It can be proved similarly that $S(A, C) \vee S(B, C) \geq S(A \wedge B, C)$.

S8 This follows from **S5**, using $C = D$ and noting that $S(A, B) \wedge S(C, C) = S(A, B) \wedge \underline{1} = S(A, B)$.

S9 Choose any $u \in U$ and consider two cases. (a) If $S_u(A, A \vee B) = 1$, then $A_u = A_u \vee B_u \Rightarrow A_u \geq B_u \Rightarrow S_u(B, A \wedge B) = 1$. (b) If $S_u(A, A \vee B) = 0$, then $A_u \neq A_u \vee B_u \Rightarrow A_u < B_u \Rightarrow A_u \wedge B_u = A_u < B_u \Rightarrow \ S_u(B, A \wedge B) = 0$. Hence, for all $u \in U$ we have $S_u(A, A \vee B) = S_u(B, A \wedge B)$, i.e. $S(A, A \vee B) = S(B, A \wedge B)$. By interchanging the role of A and B, we obtain $S(B, A \vee B) = S(A, A \wedge B)$.

S10 We will prove this by showing that (i) is equal to each of (ii) – (vi).

(i)=(ii) From symmetry and transitivity we have

$$S(A, B) = S(B, A) \geq S(B, A\vee B)\wedge S(A\vee B, A) \Rightarrow$$

(by use of **S9b** and symmetry)

$$S(A, B) \geq S(A, A\wedge B)\wedge S(A, A\vee B). \tag{13}$$

Also, if in **S6** we substitute B with A and C with B, we obtain (using $S(A, A) \wedge S(A, B) =$

$\underline{1} \wedge S(A,B) = S(A,B))$

$$S(A,A) \wedge S(A,B) \leq \left\{ \begin{array}{c} S(A, A \vee B) \\ S(A, A \wedge B) \end{array} \right\} \Rightarrow$$

$$S(A,B) \leq S(A, A \wedge B) \wedge S(A, A \vee B). \tag{14}$$

From (13,14) we see that $S(A,B) = S(A, A \wedge B) \wedge S(A, A \vee B) = S(A \wedge B, A) \wedge S(A, A \vee B)$.

(i)=(iii) This is proved exactly as above, interchanging the role of A and B.

(i)=(iv) We have $A \wedge B \leq A \leq A \vee B$. Using **S4** and that (i)=(ii), we immediately get $S(A \wedge B, A \vee B) = S(A \wedge B, A) \wedge S(A, A \vee B) = S(A,B)$.

(i)=(v) We have (using **S9b** and symmetry)

$$S(A, A \vee B) \wedge S(B, A \vee B) = S(A, A \vee B) \wedge S(A, A \wedge B) = S(A, A \vee B) \wedge S(A \wedge B, A) \tag{15}$$

but we have already proved that $S(A,B) = S(A \wedge B, A) \wedge S(A, A \vee B)$ and so the proof is complete.

(i)=(vi) This proved similarly to the previous step.

From the last few steps we see that (i) = (ii) = ... = (vi). This completes the proof of **S10** and of the proposition.■

Remark. Property **S4** is related to the concepts of *betweenness* and *convexity*; property **S10** is related to modularity. However, since these concepts are usually related to distance, rather than similarity, we will present the corresponding remarks in Section 4.2.

Proposition 4.6 *For all $A, B, C \in \mathbf{F}$ and all $\Theta \in \mathbf{B}$ we have:*

1. $S(A,B) \geq \Theta \Rightarrow S(A \vee C, B \vee C) \geq \Theta$.

2. $S(A,B) \geq \Theta \Rightarrow S(A \wedge C, B \wedge C) \geq \Theta$.

3. $S(A \wedge C, B \wedge C) = S(A \vee C, B \vee C) = \Theta \Rightarrow S(A,B) = \Theta$.

Proof. Choose any $A, B, C \in \mathbf{F}$ and any $\Theta \in \mathbf{B}$. Then we have the following.

1. This follows from **S8**.

2. This follows from **S8**.

3. Take any $A, B, C \in \mathbf{F}$ and any $\Theta \in \mathbf{B}$ such that $S(A \wedge C, B \wedge C) = S(A \vee C, B \vee C) = \Theta$. Choose any $u \in U$. We consider two cases. (a) If $\Theta_u = 1$, then $A_u \wedge C_u = B_u \wedge C_u$ and $A_u \vee C_u = B_u \vee C_u$. Then, by distributivity we have $A_u = B_u \Rightarrow S_u(A, B) = 1 = \Theta_u$. (b) If $\Theta_u = 0$, then $A_u \wedge C_u \neq B_u \wedge C_u$ and $A_u \vee C_u \neq B_u \vee C_u$ and so, obviously, $A_u \neq B_u \Rightarrow S_u(A, B) = 0 = \Theta_u$. Hence, for all $u \in U$ we have $S_u(A, B) = \Theta_u$.■

Remark. Proposition 4.6 is related to the concept of ϵ-similarity introduced by Pappis [21, 22, 23].

4.2 L-Fuzzy Distance

Definition 4.7 *For $A, B \in \mathbf{F}$, the distance between A and B is denoted by $D(A, B)$ and is defined by*

$$D(A, B) \doteq S'(A, B).$$

Example 4.8 *Continuing with the sets of the previous examples, we have: $D(\Phi, \Theta) = S'(\Phi, \Theta) = [0, 1, 0, 1]' = [1, 0, 1, 0]$; and $D(\Phi, A) = S'(\Phi, A) = [0, 0, 1, 0]' = [1, 1, 0, 1]$.*

Proposition 4.9 *For all $A, B \in \mathbf{F}$ and for all $u \in U$ we have: $D_u(A, B) = 0 \Leftrightarrow A_u = B_u$.*

Proof. Take any $A, B \in \mathbf{F}$ and any $u \in U$. $D_u(A, B) = 0 \Leftrightarrow S_u(A, B) = 1 \Leftrightarrow A_u = B_u$.■

As can be seen by the definition and by the above proposition, distance is defined as the complement of similarity. Given that complementation is order-inverting, it follows that if sets A and B have large similarity, then they will have small distance. Since distance is usually perceived as a totally ordered, nonnegative quantity, the above definition may appear rather unusual. In fact however, $D(.,.)$ has the basic characteristics of a distance function, as outlined in the next proposition.

Proposition 4.10 *For all $A, B, C \in \mathbf{F}$ we have:*

1. $D(A,B) = \underline{0} \Leftrightarrow A = B$.

2. $D(A,B) = D(B,A)$.

3. $D(A,B) \leq D(A,C) \vee D(C,B)$.

Proof. This follows from the definition $D(A,B) = S'(A,B)$ and from Proposition 4.4.■

Remark. $D(A,B) \leq D(A,C) \vee D(C,B)$ is the partial order analog of the triangle inequality (in addition it is the *ultrametric* triangle inequality). The idea of Boolean valued distances is not new; it appears already in the 1950's (for instance see Blumenthal's book [3]). This idea has been applied to Boolean lattices (which are called "*auto*metrized" spaces since the domain of the metric distance function is the same space on which the metric is imposed). Formally, in [3] we have $D : \mathbf{B} \times \mathbf{B} \rightarrow \mathbf{B}$. In our case we have $D : \mathbf{F} \times \mathbf{F} \rightarrow \mathbf{B} \subseteq \mathbf{F}$; in this sense $(\mathbf{F}, D)$ is an autometrized space. The following proposition is also of interest.

Proposition 4.11 *For all* $\Theta, \Phi \in \mathbf{B}$ *we have* $D(\Theta, \Phi) = (\Theta' \wedge \Phi) \vee (\Theta \wedge \Phi')$.

Proof. Take any $\Theta, \Phi \in \mathbf{B}$ and any $u \in U$. By considering the four cases (a) $\Theta_u = 0$, $\Phi_u = 0$, (b) $\Theta_u = 0$, $\Phi_u = 1$, (c) $\Theta_u = 1$, $\Phi_u = 0$, (d) $\Theta_u = 1$, $\Phi_u = 1$, we see that in every case (i.e. for all $u \in U$) we have $D_u(\Theta, \Phi) = (\Theta'_u \wedge \Phi_u) \vee (\Theta_u \wedge \Phi'_u)$ and the proof is complete.■

Let us present some further properties of $D(.,.)$.

Proposition 4.12 *For all* $A, B, C, E \in \mathbf{F}$ *we have:*

D1 $D(A,B) = \underline{0} \Leftrightarrow A = B$.

D2a $D(A,A') = \underline{1} \Leftrightarrow (\forall u \in U$ *we have* $A_u \neq 1/2)$.

D2b $D(A,A') = \underline{0} \Leftrightarrow (\forall u \in U$ *we have* $A_u = 1/2)$.

D3 $D(A,B) = D(A',B')$.

D4 $A \leq B \leq C \Rightarrow \begin{cases} D(A,B) \leq D(A,C) \\ D(B,C) \leq D(A,C) \\ D(A,C) = D(A,B) \vee D(B,C) \end{cases}$.

D5 $D(A,B) \vee D(C,E) \geq \begin{cases} D(A \vee C, B \vee E) \\ D(A \wedge C, B \wedge E) \end{cases}.$

D6 $D(A,B) \vee D(A,C) \geq \begin{cases} D(A, B \vee C) \\ D(A, B \wedge C) \end{cases}.$

D7 $D(A,C) \wedge D(B,C) \leq \begin{cases} D(A \vee B, C) \\ D(A \wedge B, C) \end{cases}.$

D8 $D(A,B) \geq \begin{cases} D(A \vee C, B \vee C) \\ D(A \wedge C, B \wedge C) \end{cases}$

D9a $D(A, A \vee B) = D(B, A \wedge B)$.

D9b $D(A, A \wedge B) = D(B, A \vee B)$.

D10 *All of the following quantities are equal:*

$$\begin{array}{llll} (i) & D(A,B) & (iv) & D(A \wedge B, A \vee B) \\ (ii) & D(A \wedge B, A) \vee D(A, A \vee B) & (v) & D(A, A \vee B) \vee D(B, A \vee B) \\ (iii) & D(A \wedge B, B) \vee D(B, A \vee B) & (vi) & D(A, A \wedge B) \vee D(B, A \wedge B) \end{array}.$$

Proof. This follows from the definition of $D(A,B) \doteq S'(A,B)$ and Proposition 4.5.■

Remark. Property **D4** is related to the concepts of *betweenness* and *convexity*. Consider first an abstract metric space ($\mathbf{X}$,d), where $d : \mathbf{X} \times \mathbf{X} \to [0, \infty)$ is a scalar distance function. A point b is said to be (on a *straight line segment*) *between* points $a, c \in \mathbf{X}$ iff $d(a,b) + d(b,c) = d(a,b)$ (this is in direct analogy to the case of a Euclidean space). Now, in complete analogy, in the autometrized space ($\mathbf{F}, D$), a "point" (actually fuzzy set) C is said to be between $A, B \in \mathbf{F}$ iff $D(A,C) = D(A,B) \vee D(B,C)$. It can be shown [3] that these C are exactly the ones which also satisfy $A \wedge C \leq B \leq A \vee C$. In short, "points" which are *order*-between A and C are also *metrically*-between A and C. The condition $A \leq B \leq C$, appearing in Property **D4**, is a special case of the condition $A \wedge C \leq B \leq A \vee C$. Convexity can be defined as follows: a collection of fuzzy sets $\mathbf{W} \subseteq \mathbf{F}$ is (metric- or order-) convex iff for any $A, C \in \mathbf{W}$

we have $D(A, C) = D(A, B) \vee D(B, C) \Rightarrow B \in \mathbf{W}$ (i.e. every point between A and C is contained in $\mathbf{W}$).

Remark. Similarly, Property **D10** is related to lattice modularity. Take a modular lattice $(\mathbf{X}, \leq)$ with a *positive valuation* $v(.)$ and define a distance $d(.,.)$ by $d(x, y) = v(x \vee y) - v(x \wedge y)$ (for details see [1]). Then it is easy to prove $d(x \vee y, x \wedge y) = d(x, y)$, which is analogous to the equality of (i) and (iv) in **D10**.

Proposition 4.13 *For all $A, B, C \in \mathbf{F}$ and all $\Theta \in \mathbf{B}$ we have:*

1. $D(A, B) \leq \Theta \Rightarrow D(A \vee C, B \vee C) \leq \Theta$.

2. $D(A, B) \leq \Theta \Rightarrow D(A \wedge C, B \wedge C) \leq \Theta$.

3. $D(A \wedge C, B \wedge C) = D(A \vee C, B \vee C) = \Theta \Rightarrow D(A, B) = \Theta$.

Proof. This follows from the definition of $D(A, B) \doteq S'(A, B)$ and Proposition 4.6.■

5 Conclusion

We have introduced a novel L-fuzzy valued measure of inclusion $I(.,.)$, and have established a number of its properties. Some of these properties are analogous to the ones usually postulated for "scalar" fuzzy inclusion measures. In addition, we find it particularly attractive that $I(.,.)$ turns out to be a fuzzy order. Furthermore, $I(.,.)$ can be used to define a L-fuzzy similarity and a L-fuzzy distance between fuzzy sets. Let us conclude by discussing some future research directions.

L-fuzzy order and lattice. As already mentioned, $I(.,.)$ is a L-fuzzy order relationship. In fact this order relationship can be denoted in an alternative, more suggestive manner. Rather than writing $I(A, B) = \Theta$ (where $\Theta \in \mathbf{B}$) we can also write $A \leq_\Theta B$. Now, it is easy to show that for every fixed value of Θ, $\leq_\Theta$ is a *crisp preorder* on $\mathbf{F}$. The following problem arises naturally.

Problem 5.1 *Define a L-fuzzy lattice in such a manner that the family $\{(\mathbf{F}, \leq_\Theta)\}_{\Theta \in \mathbf{B}}$ are its* cuts. *Develop the corresponding L-fuzzy lattice theory.*

Implication. The relationship between set inclusion and logical implication is a well known one (as discussed in several parts of this paper). Our L-fuzzy inclusion measure can be viewed as an *L-fuzzy valued implication operators*. This connection will be reported elsewhere and parallels the work of many authors in obtaining inclusion measures from fuzzy implication operators, in the style of Bandler and Kohout. The L-fuzzy implication operator satisfies properties analogous to Klir's axioms for the implication operator [15]. The connection to *conditional probability* is also worth investigating. A scalar inclusion measure $i(A, B)$ is in many ways analogous to $\Pr(B|A)$ (compare with inclusion measure no.1 in Section 3.2); it would be interesting to use the ideas presented in this paper to define lattice-valued probability measures.

Aggregation. It is interesting to investigate the existence of other "vector" inclusion measures. In particular, "scalar" inclusion measures and our $I(.,.)$ lie at extreme ends of a spectrum. A scalar inclusion measure aggregates the inclusion information of *all* elements of a set into a single "global" value; our $I(.,.)$ preserves all the "local" information about elementwise inclusion. Perhaps inclusion measures which are halfway between extreme aggregation and extreme localization will also prove useful. A route to arrive at such inclusion measures could be the following: define a partition $\{U_1, U_2, ..., U_K\}$ of U and then define on $\mathbf{F} \times \mathbf{F}$ a vector inclusion measure $\widetilde{I}(A, B) \doteq [\widetilde{i}_1(A, B), \widetilde{i}_2(A, B), \ldots, \widetilde{i}_K(A, B)]$ where, for $k = 1, 2, ..., K$, $\widetilde{i}_k(A, B)$ is a scalar inclusion measure which depends only on values of A_u, B_u, for $u \in U_k$. In this manner, the details about elementwise inclusion are not aggregated into a single number; some degree of local information is preserved.

In this connection, let us mention the approach to *similarity* measures appearing in [12]. In this paper, the authors view a similarity measure as the relation resulting from similarity between degrees of membership as *local* relations. They relate this approach to implication operators and discuss transitivity. While in the above paper the local *preferences* are aggregated (unlike our own approach) we find interesting the explicit recognition that similarity (and inclusion) is determined in terms of many local relations. This is exactly the approach we are taking in this paper, except that we do not perform the aggregation step.

References

[1] G. Birkhoff, *Lattice Theory*, American Mathematical Society, Colloquium Publications, vol. 25, 1967.

[2] W. Bandler and L. Kohout. "Fuzzy power sets and fuzzy implication operators". *Fuzzy Sets and Systems*, vol.4, pp.13–30, 1980.

[3] L.M. Blumenthal. *Theory and Applications of Distance Geometry*. Oxford, Clarendon Press, 1953.

[4] S. Bodjanova. "Approximation of fuzzy concepts in decision making". *Fuzzy Sets and Systems*, vol.85, pp.23–29, 1997.

[5] B. Bouchon-Meunier, M. Rifqi and S. Bothorel. "Towards general measures of comparison of objects". *Fuzzy Sets and Systems*, vol. 84, pp. 143–153, 1996.

[6] P. Burillo, N. Frago and R. Fuentes. "Inclusion grade and fuzzy implication operators". *Fuzzy Sets and Systems*, vol.114, pp. 417–429, 2000.

[7] H. Bustince. "Indicator of inclusion grade for interval-valued fuzzy sets. Application to approximate reasoning based on interval-valued fuzzy sets". *Int. J. of Approximate Reasoning*, vol.23, pp.137–209, 2000.

[8] B.A. Davey and H.A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, 1990.

[9] J. Fan, W. Xie and J. Pei. "Subsethood measure: new definitions". *Fuzzy Sets and Systems*, vol.106, pp.201–209, 1999.

[10] J. Fan and W.Xie. "Some notes on similarity measure and proximity measure". *Fuzzy Sets and Systems*, vol.101, pp.403–412, 1999.

[11] J.C. Fodor and S. Ovchinnikov. "On aggregation of T-transitive fuzzy binary relations". *Fuzzy Sets and Systems*, vol.72, pp.135–145, 1995.

[12] P. Fonck, J. Fodor and M. Roubens. "An application of aggregation procedures to the definition of measures of similarity between fuzzy sets". *Fuzzy Sets and Systems*, vol.97, pp.67–74, 1998.

[13] V.G. Kaburlasos and V. Petridis. "Learning and decision-making in the framework of fuzzy lattices", in *New Learning Techniques in Computational Intelligence Paradigms*, ed. L.C. Jain. Boca Raton, FL: CRC Press, 2000.

[14] V.G. Kaburlasos and V. Petridis. "Fuzzy lattice neurocomputing (FLN) : a novel connectionist scheme for versatile learning and decision daking by clustering", *International Journal of Computers and Their Applications*, vol. 4, pp. 31-43, 1997.

[15] G.J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, 1991.

[16] B. Kosko. *Neural Networks and Fuzzy Systems*. Prentice-Hall, 1992.

[17] S. Kundu. "A representation theorem for min-transitive fuzzy relations". *Fuzzy Sets and Systems*, vol.109, pp.453–457, 2000.

[18] X.C. Liu, "Entropy, distance measure and similarity measure of fuzzy sets and their relations", *Fuzzy Sets and Systems*, vol.52 , pp. 305–318, 1992.

[19] H.T. Ngyen and E.A. Walker. *A First Course in Fuzzy Logic*. CRC Press 1997.

[20] S. Ovchinnikov. "Similarity relation, fuzzy partitions and fuzzy orderings". *Fuzzy Sets and Systems*, vol.40, pp.107–126, 1991.

[21] C.P. Pappis. "Value approximation of fuzzy systems variables". *Fuzzy Sets and Systems*, vol.39, pp.111–115, 1991.

[22] C.P. Pappis and N.I. Karacapilidis. "A comparative assessment of measures of similarity of fuzzy values." *Fuzzy Sets and Systems*, vol.56, pp.171–174, 1993.

[23] C.P. Pappis and N.I. Karacapilidis. "Application of a similarity measure of fuzzy sets to fuzzy relational equations". *Fuzzy Sets and Systems*, vol.75, pp.135–142, 1992.

[24] V. Petridis and V.G. Kaburlasos. "Fuzzy lattice neural network (FLNN): a hybrid model for learning, *IEEE Transactions on Neural Networks*, vol. 9, p. 877-890, 1998.

[25] B. Seselja and A. Tepavcevic. "Partially ordered and relational valued fuzzy relations I". *Fuzzy Sets and Systems*, vol.72, pp.205–213, 1995.

[26] D. Sinha and E.R. Dougherty. "Fuzzification of set inclusion: theory and applications". *Fuzzy Sets and Systems*, vol. 55, pp. 15–42, 1993.

[27] A. Tepavcevic and G. Trajkovski, "Introduction to lattice valued fuzzy lattices", to appear in *Fuzzy Sets and Systems.*

[28] W-J. Wang. "New similarity measures on fuzzy sets and elements". *Fuzzy Sets and Systems*, vol.85, pp.305–309, 1997.

[29] X.Wang, B. deBaets and E. Kerre. "A comparative study of similarity measures". *Fuzzy Sets and Systems*, vol.73, pp.259–268, 1995.

[30] R. Willmott, "On the transitivity of containment and equivalence in fuzzy power set theory", *J. Math. Anal. and Appl.*, vol. 120, pp.384-396, 1986.

[31] R. Willmott, "Two fuzzier implication operators in the theory of fuzzy power sets", *Fuzzy sets and Systems*, vol.4, pp.31-36, 1980.

[32] V. R. Young. "Fuzzy subsethood". *Fuzzy Sets and Systems*, vol.77, pp.371–384, 1996.

[33] L.A. Zadeh, "Fuzzy Sets", *Information and Control*, vol.8, pp.338-353, 1965.