Inclusion Measures in Intuitionistic Fuzzy Set Theory

Chris Cornelis and Etienne Kerre

Department of Mathematics and Computer Science, Fuzziness and Uncertainty Modelling Research Unit, Ghent University, Krijgslaan 281 (S9), B-9000 Gent, Belgium {chris.cornelis, etienne.kerre}@rug.ac.be
http://fuzzy.rug.ac.be

Abstract. Twenty years after their inception, intuitionistic fuzzy sets are on the rise towards making their "claim to fame". Competing alongside various other, often closely related, formalisms, they are catering to the needs of a more demanding and rapidly expanding knowledge–based systems industry. In this paper, we develop the notion of a graded inclusion indicator within this setting, drawing inspiration from related concepts in fuzzy set theory, yet keeping a keen eye on those particular challenges raised specifically by intuitionistic fuzzy set theory. The use of our work is demonstrated by its applications in approximate reasoning and non–probabilistic entropy calculation.

1 Introduction and Problem Definition

1.1 Putting Intuitionistic Fuzzy Set Theory on the Map

IFS theory basically enriches Zadeh's fuzzy set theory with a notion of indeterminacy expressing hesitation or abstention. While in the latter, membership degrees, identifying the degree to which an object satisfies a given property (generally speaking), are taken to be exact, in the former extra information in the guise of a non–membership degree is permitted to address a commonplace feature of uncertainty. Imagine, for instance, a voting procedure in which delegates have to express their feelings w.r.t. a number of proposals. It is obvious that while one can be in favour or in disfavour of a proposal to a certain extent, one can also abstain from the vote; an attitude inspired by, e.g., a lack of background or interest, or simply because no obvious arguments for or against the cause at stake have been raised. In such a situation, using only a $[0,1]$–valued degree α expressing support for the proposal is arguably too committing. A similar argument can be set up when the opinion of a given voter is not (fully) known, and we should be duly hesitant to classify him as a supporter or an opponent of the proposal.

IFS theory allows for an easy, yet elegant, way out of such problems by not insisting that membership and non–membership to a set be strictly complementary properties. In an IFS A defined in a universe[1] X, alongside a **membership**

[1] For simplicity, throughout this paper X is assumed to be finite.

degree $\mu_A(x)$ of x to A, we also distinguish a non–membership degree $\nu_A(x)$, such that $\mu_A(x) + \nu_A(x) \leq 1$. Note that a fuzzy set in X is then just an IFS for which $\mu_A(x)+\nu_A(x) = 1$ holds for every x. The degree $\pi_A(x) = 1-\mu_A(x)-\nu_A(x)$ quantifies the degree of indeterminacy associated with x and A.

Just like the relationship between classical logic and set theory was exploited in fuzzy set theory to define "fuzzy logics" (in a narrow sense), so we may also introduce a notion of "intuitionistic fuzzy (IF) logics"; with a proposition P a degree of truth μ_P and one of falsity ν_P may be associated, such that $\mu_P+\nu_P \leq 1$. This idea is elaborated in e.g. [1].

As it turns out, IFSs pop up quite naturally. Attempts to embed IFS theory within more "familiar" frameworks have shown that they fit in with, and enrich, a well–established tradition of modeling imprecision rather than setting off on an entirely new course, which marks their relevance. It can easily be seen, for instance, that IFSs are formally equivalent to interval–valued fuzzy sets: indeed, a couple $(\mu_A(x), \nu_A(x))$ may be mapped bijectively onto an interval $[\mu_A(u), 1 - \nu_A(u)]$. Some would consider this syntactical equivalence sufficient evidence to dismiss IFS theory as superfluous and giving cause to unnecessary confusion. We raise two arguments against such allegations:

1. Interval–valued fuzzy set theory is currently associated, *de facto*, with the work of Mendel and others on type–2 fuzzy logic systems (see e.g. [13]). That setting is characterized by a probability–like treatment of uncertainty on the membership degrees in a fuzzy set; an interval–valued fuzzy set is designated as a special type–2 fuzzy set[2] that exhibits a uniform spread of uncertainty on the membership degrees. IFS theory, however, does not make any assumptions on the nature of its indeterminacy—it merely gives a quantitative representation of "missing information".
2. We consider IFS theory as a stepping stone in a larger context that is specifically tuned to the concept of positive and negative constituents, rather than lower and upper approximations. Indeed, if we relax the constraint that $\mu_A(x)$ and $\nu_A(x)$ sum up to at most 1, letting either degree range freely in $[0, 1]$, we arrive in the realm of fuzzy four–valued logics first suggested by Stickel [15] and given a nice practical application by Fortemps and Slowinski [9], who used degrees $(\alpha, \beta) \in [0, 1]^2$ whose respective components express positive and negative evidence in a preference setting. It is clear that as soon as $\alpha + \beta > 1$, evidence is inconsistent to some extent. In that sense, IFS theory can be seen as the consistent restriction of the fuzzy four–valued framework.

Unfortunately, also a lot of misunderstandings concerning terminology have sprung up. The term "intuitionistic" is to be read in a "broad" sense here, alluding loosely to the denial of the law of the excluded middle on element level (since $\mu_A(x) + \nu_A(x) < 1$ is possible). A "narrow", graded extension of intuitionistic logic proper has also been proposed and is due to Takeuti and Titani [17]—it bears no relationship to our notion of IFS theory.

[2] i.e. a fuzzy set whose membership degrees are themselves fuzzy sets in $[0, 1]$

1.2 An Introduction to Graded Inclusion Measures

In fuzzy set theory, inclusion is, by default, defined as follows: for A and B fuzzy sets[3] in a universe X, $A \subseteq B \iff (\forall x \in X)(A(x) \leq B(x))$, i.e. $A \subseteq B$ if and only if the graph of A fits beneath the graph of B. A natural extension of this definition to IFS theory reads, for A and B IFSs in X: $A \subseteq B \iff \mu_A \subseteq \mu_B$ and $\nu_B \subseteq \nu_A$.

While in many theoretical and practical settings this two–valued characterization of subsethood suffices, it could be argued that the definition is overly restrictive: just as an element can belong to a fuzzy set to varying degrees, so we may also want to talk about a fuzzy set being "more or less" a subset of another one. Many researchers [2, 7, 8, 11, 12, 14, 18] have tried to capture this intuition by proposing concrete operators Inc that take a couple of fuzzy sets (A, B) as their input and return a value $Inc(A, B)$ in $[0, 1]$ indicating the degree of subsethood of A to B.

Typically, to define fuzzy subsethood one takes a definition of classical set inclusion and tries to extend ("fuzzify") it to apply to fuzzy sets. Below we quote three distinct, but essentially equivalent[4], definitions of the inclusion of A into B, where $A, B \in \mathcal{P}(X)$:

$$A \subseteq B \iff (\forall x \in X)(x \in A \Rightarrow x \in B), \tag{1}$$

$$\iff A = \emptyset \text{ or } \frac{|A \cap B|}{|A|} = 1, \tag{2}$$

$$\iff \frac{|co(A) \cup B|}{|X|} = 1 \tag{3}$$

While (1) is stated in strictly logical terms, the other two are based on counting the elements of a set, i.e. on cardinality, and have a probabilistic (i.e. frequentist) touch about them. It is therefore not surprising that their respective generalizations to fuzzy set theory cease to be equivalent. Without going into the details at this point, we might roughly state that adepts of the different crisp definitions have put fuzzy subsethood on two separate tracks, one logic–based, the other frequency–based. One situation where this distinction comes to light is when one tries to mould fuzzy inclusion measures into axiomatic characterizations by listing desirable properties for them, as several authors have attempted. The most strident dissonance (see e.g. Young [18] on this) seems to concern the condition

$$A, B \in \mathcal{P}(X) \Rightarrow Inc(A, B) \in \{0, 1\} \tag{4}$$

called heritage by Kitainik [11]. As will be revealed later on, choosing to impose it pretty much forces us into the logic–based approach, although useful trade–offs are possible.

[3] For simplicity, we identify a fuzzy set A with its membership function μ_A and write $A(x)$ to denote $\mu_A(x)$.

[4] Arguably, (1) is more general since it can also deal with infinite sets.

In this paper, we are going to pursue this discussion to the framework of IFS theory. Our aim is twofold: first we are going to try and convey as complete and uniform as possible a picture of IF inclusion by contrasting and generalizing corresponding fuzzy approaches; and secondly, we will highlight a few distinguishing features that are specific only to the extension, and which are meant to refute the criticism that IF subsethood assessment merely amounts to applying fuzzy inclusion measures twice.

The paper is organized as follows: in section 2, we recall the necessary mathematical background on IFS theory. Section 3 starts by investigating what an IF inclusion measure should look like, and what properties it should ideally satisfy. This results in the development of logic– and frequency–based approaches. In section 4, we briefly sketch the application of these measures in two concrete domains: approximate reasoning and entropy measurement. Finally, section 5 offers a brief conclusion.

2 Preliminaries of Intuitionistic Fuzzy Set Theory

Atanassov [1] gives the following definition of an IFS A in X:

$$A = \{(x, \mu_A(x), \nu_A(x)) \mid x \in X\} \tag{5}$$

where μ_A and ν_A are called membership and non–membership function of A respectively. They satisfy $\mu_A(x) + \nu_A(x) \leq 1$ for every $x \in X$. The class of all IFSs in X is denoted $\mathcal{IF}(X)$.

This definition is easy to absorb for humans but lacks mathematical conciseness. Just as a fuzzy set in X can be interpreted as a mapping from X to $[0,1]$, so we may define an IFS A in X as a mapping from X to the set $L^* = \{(x_1, x_2) \in [0,1]^2 \mid x_1 + x_2 \leq 1\}$. Moreover, equiping L^* with an ordering $\leq_{L^*}$ defined as $(x_1, x_2) \leq_{L^*} (y_1, y_2) \Leftrightarrow x_1 \leq y_1$ and $x_2 \geq y_2$, $(L^*, \leq_{L^*})$ assumes the structure of a complete, bounded lattice with greatest element $1_{L^*} = (1,0)$ and smallest element $0_{L^*} = (0,1)$. The sup and inf operations on this lattice are derived from $\leq_{L^*}$ as:

$$\sup((x_1, y_1), (x_2, y_2)) = (\max(x_1, x_2), \min(y_1, y_2)) \tag{6}$$

$$\inf((x_1, y_1), (x_2, y_2)) = (\min(x_1, x_2), \max(y_1, y_2)) \tag{7}$$

The intersection, union and complement of IFSs A and B in $\mathcal{IF}(X)$ are defined by, for $x = (x_1, x_2) \in L^*$, $A \cap B(x) = \inf(A(x), B(x))$, $A \cup B(x) = \sup(A(x), B(x))$, $co(A)(x) = A(x_2, x_1)$. Thus, IFSs are a special case of L–fuzzy sets in the sense of Goguen [10], with $L = L^*$. As a shorthand notation, for $x \in L^*$, we denote its first, resp. second component by x_1 and x_2. A special subset D of "fuzzy values" of L^* is defined by $D = \{(x_1, x_2) \in L^* \mid x_1 = 1 - x_2\}$.

Since $\leq_{L^*}$ is a partial order, an order–theoretic extension of classical negation, conjunction, disjunction and implication on L^*, as negators, triangular norms and conorms, and implicators, respectively, arises quite naturally: a negator on L^* is any decreasing $L^* \to L^*$ mapping $\mathcal{N}$ that satisfies $\mathcal{N}(0_{L^*}) = 1_{L^*}$ and

$\mathcal{N}(1_{L^*}) = 0_{L^*}$. The mapping $\mathcal{N}_s$, defined as $\mathcal{N}_s(x_1, x_2) = (x_2, x_1), \forall (x_1, x_2) \in L^*$, will be called the *standard negator.*

A t–norm on L^* is any increasing, commutative, associative $(L^*)^2 \to L^*$ mapping $\mathcal{T}$ that satisfies $\mathcal{T}(1_{L^*}, x) = x$, for all $x \in L^*$; a t–conorm on L^* is any increasing, commutative, associative $(L^*)^2 \to L^*$ mapping $\mathcal{S}$ satisfying $\mathcal{S}(0_{L^*}, x) = x$, for all $x \in L^*$. Obviously, the greatest t–norm with respect to the ordering $\leq_{L^*}$ is inf, while the smallest t–conorm w.r.t. $\leq_{L^*}$ is sup. Note that it does not hold that for all $x, y \in L^*$, either $\inf(x, y) = x$ or $\inf(x, y) = y$. For instance, $\inf((0.1, 0.3), (0.2, 0.4)) = (0.1, 0.4)$. t–norms and t–conorms can be partitioned into two classes by the following definition: a t–norm $\mathcal{T}$ on L^* (resp. t–conorm $\mathcal{S}$) is called t–representable if there exists a t–norm T and a t–conorm S on $[0, 1]$ (resp. a t–conorm S' and a t–norm T' on $[0, 1]$) such that, for $x = (x_1, x_2), y = (y_1, y_2) \in L^*$, $\mathcal{T}(x, y) = (T(x_1, y_1), S(x_2, y_2)), \mathcal{S}(x, y) = (S'(x_1, y_1), T'(x_2, y_2))$; T and S (resp. S' and T') are called the representants of $\mathcal{T}$ (resp. $\mathcal{S}$). Clearly, inf and sup are t–representable. The following mappings $\mathcal{T}_W$ and $\mathcal{S}_W$, called IF Łukasiewicz t–norm and t–conorm, are not [4]:

$$\mathcal{T}_W(x, y) = (\max(0, x_1 + y_1 - 1), \min(1, x_2 + 1 - y_1, y_2 + 1 - x_1)) \tag{8}$$
$$\mathcal{S}_W(x, y) = (\min(1, x_1 + 1 - y_2, y_1 + 1 - x_2), \max(0, x_2 + y_2 - 1)) \tag{9}$$

It can be verified that $\mathcal{T}_W$ is a t–norm on L^* and $\mathcal{S}_W$ a t–conorm on L^*; their existence rules out the conjecture, implicit in most of the existing literature that t–norms and t–conorms on L^* are necessarily characterized by a pair of fuzzy connectives.

Negators, t–norms and t–conorms on L^* may be used to define generalized versions of complementation, intersection and union of IFSs. Specifically, we may define $co_{\mathcal{N}}(A)$, $A \cap_{\mathcal{T}} B$ and $A \cup_{\mathcal{T}} B$ by $co_{\mathcal{N}}(A)(x) = \mathcal{N}(A(x))$, $A \cap_{\mathcal{T}} B(x) = \mathcal{T}(A(x), B(x))$, $A \cup_{\mathcal{S}} B(x) = \mathcal{S}(A(x), B(x))$, where $x \in X$.

The final and for our purposes most important construct is that of an implicator on L^*: an $(L^*)^2 \to L^*$–mapping $\mathcal{I}$ satisfying $\mathcal{I}(0_{L^*}, 0_{L^*}) = 1_{L^*}, \mathcal{I}(1_{L^*}, 0_{L^*}) = 0_{L^*}, \mathcal{I}(0_{L^*}, 1_{L^*}) = 1_{L^*}, \mathcal{I}(1_{L^*}, 1_{L^*}) = 1_{L^*}$. Moreover we require $\mathcal{I}$ to be decreasing in its first, and increasing in its second component. This definition is very general; as in fuzzy set theory, we may distinguish implicators on L^* w.r.t. their construction. Explicitly, an S–implicator $\mathcal{I}_{\mathcal{S},\mathcal{N}}$ is defined as, for $x, y \in L^*$:

$$\mathcal{I}_{\mathcal{S},\mathcal{N}}(x, y) = \mathcal{S}(\mathcal{N}(x), y) \tag{10}$$

with $\mathcal{S}$ a t–conorm and $\mathcal{N}$ a negator on L^*. An R–implicator $\mathcal{I}_{\mathcal{T}}$, generated by a t–norm $\mathcal{T}$ on L^* is defined as, for $x, y \in L^*$, by

$$\mathcal{I}_{\mathcal{T}}(x, y) = \sup\{\gamma \in L^* \mid \mathcal{T}(x, \gamma) \leq_{L^*} y\} \tag{11}$$

These two classes contain most of the prominent implicators. For example, the S–implicator of $\mathcal{S}_W$ and $\mathcal{N}_s$, equal to the R–implicator of $\mathcal{T}_W$ is given by

$$\mathcal{I}_{\mathcal{S}_W,\mathcal{N}_s} = \mathcal{I}_{\mathcal{T}_W}(x, y) = (\min(1, y_1 + 1 - x_1, x_2 + 1 - y_2), \max(0, y_2 + x_1 - 1)) \tag{12}$$

Other than by their construction, implicators on L^* may also be classified by the properties they satisfy. The following important theorem is based on [4].

Theorem 1 *A continuous implicator $\mathcal{I}$ on L^* satisfies*

$$(\forall x, y, z \in L^*)(\mathcal{I}(x, \inf(y, z)) = \inf(\mathcal{I}(x, y), \mathcal{I}(x, z)))) \tag{13}$$
$$(\forall x \in L^*)(\mathcal{I}(1, x) = x) \tag{14}$$
$$(\forall x, y \in L^*)(\mathcal{I}(\mathcal{N}_s(y), \mathcal{N}_s(x)) = \mathcal{I}(x, y)) \tag{15}$$
$$(\forall x, y, z \in L^*)(\mathcal{I}(x, \mathcal{I}(y, z)) = \mathcal{I}(y, \mathcal{I}(x, z))) \tag{16}$$
$$(\forall x, y \in L^*)(x \leq_{L^*} y \iff \mathcal{I}(x, y) = 1_{L^*}) \tag{17}$$
$$(\forall x, y \in L^*)(x = 1_{L^*} \text{ and } y = 0_{L^*} \iff \mathcal{I}(x, y) = 0_{L^*}) \tag{18}$$
$$\mathcal{I}(D, D) \subseteq D \tag{19}$$

iff there exists a continuous increasing permutation[5] ϕ *of* $[0, 1]$ *s. t., for* $x, y \in L^*$,

$$\mathcal{I}(x, y) = (\varphi^{-1} \min(1, 1 + \varphi(y_1) - \varphi(x_1), 1 + \varphi(1 - y_2) - \varphi(1 - x_2)),$$
$$1 - \varphi^{-1} \min(1, 1 - \varphi(x_1) + \varphi(1 - y_2))) \tag{20}$$

To conclude this section, the cardinality of an IFS A in X was defined by Szmidt and Kacprzyk [16] as the couple $(\min \Sigma Count(A), \max \Sigma Count(A))$, where

$$\min \Sigma Count(A) = \sum_{x \in X} \mu_A(x) \tag{21}$$
$$\max \Sigma Count(A) = \sum_{x \in X} \mu_A(x) + \pi_A(x) = \sum_{x \in X} (1 - \nu_A(x)) \tag{22}$$

3 Construction of IF Inclusion Measures

In this section, we study different strategies to come up with reasonable subsethood indicators $\mathcal{I}nc$ for IFSs. A first question that needs to be answered is what kind of a mapping $\mathcal{I}nc$ should be: evidently, its inputs are IFSs in X, but what kind of object should its output be? Since we have been speaking about graded inclusion indicators, the natural option seems to be just a number in $[0, 1]$; the following example, however, shows that this strategy can lead to anomalies.

Example 1 *Let A, B be IFSs in $X = \{x_1, x_2\}$, such that $A(x_1) = 1_{L^*}$, $B(x_1) = (0, 0)$, $A(x_2) = (0, 0)$, $B(x_2) = 0_{L^*}$. Obviously, $A \not\subseteq B$. Yet, due to the indeterminacy w.r.t. B and x_1, and w.r.t. A and x_2, there is no indication that A is not a subset of B at all, nor can it be argued that A is a subset of B to a given extent $\alpha \in [0, 1]$: in fact, it* could *be a subset, to a certain extent, of B, but the presence of maximal indeterminacy does not allow to cut the knot! In this sense, forcing $\mathcal{I}nc(A, B)$ to be in $[0, 1]$ is too committing, and a more natural*

[5] It can be verified that this is equivalent to the existence of a permutation Φ of L^*, where $\Phi(x) = (\phi(x_1), 1 - \phi(1 - x_2))$, such that $\mathcal{I} = \Phi^{-1} \circ \mathcal{I}_{\mathcal{T}_W} \circ (\Phi, \Phi)$. For this reason, $\mathcal{I}$ is also called a Φ–transform of the R–implicator of the IF Łukasiewicz t–norm.

way to express A's inclusion into B is by the element $(0,0) \in L^*$*, exploiting it to express the same kind of indeterminacy as it does for partial membership to a set: we simply cannot tell.*

This example suggests that $\mathcal{I}nc$ be an $\mathcal{IF}(X) \times \mathcal{IF}(X) \to L^*$ mapping. It also presents a criterion for IF inclusion measures without an analog in fuzzy set theory, namely that $\mathcal{I}nc(A,B) = (0,0)$ when A and B are as in the above test case. On the other hand, we can borrow substantially from the available literature on fuzzy inclusion measures, as we will see shortly.

A convenient way to derive IF inclusion measures is to list a number of criteria for them, and then find out which operations satisfy these conditions. In fuzzy set theory, such an approach was taken by Sinha and Dougherty [14], and independently also by Kitainik [11]. Although neither linked their results explicitly to one of the formulas (1–3) defining crisp subsethood, subsequent research [7] pointed out that they implicitly invoked (1), and thus the use of an implicator on $[0,1]$, by their insistence on the heritage property (4). For now, we will take this property for granted; a convenient working set of criteria is then given as:

(I1) Contrapositivity	$\mathcal{I}nc(A,B) = \mathcal{I}nc(coB, coA)$
(I2) Distributivity	$\mathcal{I}nc(A, B\cap C) = \inf(\mathcal{I}nc(A,B), \mathcal{I}nc(A,C))$
(I3) Symmetry	$\mathcal{I}nc(A,B) = \mathcal{I}nc(S(A), S(B))$
(I4) Faithfulness	a) $\mathcal{I}nc(A,B) = 1_{L^*} \iff A \subseteq B$ b) $\mathcal{I}nc(A,B) = 0_{L^*} \iff$ $(\exists x \in X)(A(x) = 1_{L^*}$ and $B(x) = 0_{L^*})$ c) $A, B \in \mathcal{F}(X) \Rightarrow \mathcal{I}nc(A,B) \in D$

where $A, B, C \in \mathcal{IF}(X)$ and S an $\mathcal{IF}(X) \to \mathcal{IF}(X)$ mapping defined by, for $x \in X$, $S(A)(x) = A(s(x))$, with s a permutation of X.

Historically, these conditions go back to different sources[6]: the first three requirements were adopted from Kitainik's work on fuzzy inclusion measures [11], while the two faithfulness conditions (I4a–b) are due to Sinha and Dougherty. [14] The heritage property is a consequence of (I4a–b). Finally, we added another faithfulness condition to ensure that $\mathcal{I}nc$, when applied to fuzzy information, acts like a fuzzy inclusion measure. It can be verified that a mapping satisfying (I1–I4) is decreasing in its first, and increasing in its second component. The following theorem gives an explicit characterization. Its proof draws its inspiration from [7],[8] and [11].

Theorem 2 *An* $\mathcal{IF}(X) \times \mathcal{IF}(X) \to L^*$ *mapping* $\mathcal{I}nc$ *satisfies (I1)–(I4) iff*

$$\mathcal{I}nc(A,B) = \inf_{x\in X} \mathcal{I}(A(x), B(x)), \tag{23}$$

with $\mathcal{I}$ *an IF implicator satisfying properties (13),(15), (17), (18) and (19).*

[6] For a detailed account of the various links between Kitainik's and Sinha and Dougherty's approach, and their unification, we refer the interested reader to [7].

It is interesting that in order to be compliant with the test case of example 1, $\mathcal{I}$ should also satisfy (14). Few candidates $\mathcal{I}$ fulfill all requirements; theorem 1 characterized all continuous mappings complying with these stringent conditions, i.e. the Φ–transforms of the R–implicator of the Łukasiewicz t–norm $\mathcal{T}_W$. The simplest of these uses $\Phi(x) = x$ for all $x \in L^*$ and will be called $\mathcal{I}nc_{\mathcal{T}_W}$:

$$\mathcal{I}nc_{\mathcal{T}_W}(A, B) = \inf_{x \in X} \mathcal{I}_{\mathcal{T}_W}(A(x), B(x)) \tag{24}$$

A byproduct of this result is that it forces us to reject the argument that a graded subsethood assessment for IFSs could be reduced trivially to assessing e.g. $\mathcal{I}nc_{\mathcal{T}_W}(\mu_A, \mu_B)$ and $\mathcal{I}nc_{\mathcal{T}_W}(\nu_B, \nu_A)$, which are both in D by (I4c). Indeed, for the test case in example 1, both of these are equal to 0_{L^*}, whereas $\mathcal{I}nc_{\mathcal{T}_W}(A, B) = (0, 0)$. In other words, determining subsethood for IFSs does not amount to a mere double application of a fuzzy inclusion measure, as some IFS critics suggest!

Let us focus again on the heritage condition (4). In her paper on fuzzy subsethood, Young [18] raises skepticism about it: she reasons that much of the relative structure of fuzzy sets, and by extension IFSs, is lost when imposing it; indeed, if two fuzzy sets A and B are equal everywhere, except in the point x for which $A(x) = 1$ and $B(x) = 0$, (4) forces $Inc(A, B) = 0$. One can think of very concrete instances in which this indeed makes no sense. Imagine for instance that we are to evaluate to what extent the young people in a company are also rich. Testing subsethood of the fuzzy set of young workers into that of rich workers should then be based on the relative fraction (i.e. the *frequency*) of good earners among the youngsters, and not on whether there exists or does not exist one poor, young employee. This observation has led researchers to consider extensions to definition (2) of crisp subsethood, which works well for fuzzy inclusion measures. Indeed, if A and B are fuzzy sets, then e.g. $\min \Sigma Count(A) = \max \Sigma Count(A)$; putting $|A| = \min \Sigma Count(A)$, one can define the subsethood of A into B as the ratio of $|A \cap B|$ and $|A|$ if $A \neq \emptyset$, and 0 otherwise (see e.g. [12]). Unfortunately, there is no straightforward extension to IFS theory, since IF cardinalities are intervals of positive real values; resorting to interval calculus is not a viable option, either, as the following example shows.

Example 2 *Given two strictly positive real intervals* $[a, b]$ *and* $[c, d]$*, interval calculus defines their ratio as*

$$\frac{[a, b]}{[c, d]} = \left[\frac{a}{d}, \frac{b}{c}\right] \tag{25}$$

Define f*, for IFSs* A *and* B *in* X*, as*

$$f(A, B) = \frac{[\min \Sigma Count(A \cap B), \min \Sigma Count(A \cap B)]}{[\min \Sigma Count(A), \max \Sigma Count(A)]} \tag{26}$$

Let $X = \{x\}$*,* $A(x) = (0.5, 0.3)$*,* $B(x) = (0.6, 0.2)$*. Then* $A \subseteq B$*, but* $f(A, B) = \frac{[0.5, 0.7]}{[0.5, 0.7]} = [\frac{5}{7}, \frac{7}{5}] \neq [1, 1]$*. It is also unclear how to associate* $f(A, B)$ *with an element of* L^**.*

Evidently, the extension of (2) by interval calculus is problematic and appears to be due to definition (25). If however in (25) we put $c = d$, the result *will* be an interval in $[0, 1]$. This property is particularly useful when one considers the alternative definition (3) of crisp inclusion. Indeed, using the definition of generalized IF union and complement, we can compute

$$f_{\mathcal{S},\mathcal{N}}(A,B) = \frac{[\min \Sigma Count(co_{\mathcal{N}}(A) \cup_{\mathcal{S}} B), \max \Sigma Count(co_{\mathcal{N}}(A) \cup_{\mathcal{S}} B)]}{[\min \Sigma Count(X), \max \Sigma Count(X)]} \quad (27)$$

Let $f_{\mathcal{S},\mathcal{N}}(A,B) = [c_1, c_2]$. Since $\min \Sigma Count(X) = \max \Sigma Count(X) = |X|$, $(c_1, 1 - c_2) \in L^*$. Recalling definition (10) of an S–implicator on L^*, we may therefore introduce the following class of inclusion measures $\mathcal{I}nc_{\mathcal{S},\mathcal{N}}$:

$$\begin{aligned}\mathcal{I}nc_{\mathcal{S},\mathcal{N}}(A,B) &= \left(\frac{1}{|X|}\sum_{x\in X}(\mathcal{I}_{\mathcal{S},\mathcal{N}}(A(x),B(x)))_1, \frac{1}{|X|}\sum_{x\in X}(\mathcal{I}_{\mathcal{S},\mathcal{N}}(A(x),B(x)))_2\right)\\ &= \frac{1}{|X|}\sum_{x\in X}\mathcal{I}_{\mathcal{S},\mathcal{N}}(A(x),B(x)) \quad (28)\end{aligned}$$

The second equality introduces a convenient shorthand notation.

An example at hand is $\mathcal{I}nc_{\mathcal{S}_W,\mathcal{N}_s}$. It satisfies all of (I1–I3), (I4a) and (I4c), but not (I4b). In that sense, it has a much more lenient behaviour w.r.t. the young and rich employee problem. In fact, it can be seen that it bears a close relationship to $\mathcal{I}nc_{\mathcal{T}_W}$: instead of taking the infimum of all values $\mathcal{I}_{\mathcal{T}_W}(A(x), B(x))$, it computes their "average", which allows for a much greater deal of compensation.

4 Applications of IF Inclusion Measures

4.1 Inclusion–Based Approximate Reasoning

Roughly, approximate reasoning is concerned with the deduction of imprecise conclusions from imprecise premises. In the context of IFS theory, an IF **if–then rule** is a construct with the generic form "If $\mathcal{V}_1$ is A then $\mathcal{V}_2$ is B" where $\mathcal{V}_1$ and $\mathcal{V}_2$ represent an input and an output variable, respectively, and A and B are normalized[7] IFSs in the universes X of $\mathcal{V}_1$ and Y of $\mathcal{V}_2$. Typically, then, the system is presented with an observation on the input variable of the form "$\mathcal{V}_1$ is A'" with A' not necessarily equalling A, and asked to derive a suitable IFS B' such that "$\mathcal{V}_2$ is B'". One way of obtaining B' is by applying the Compositional Rule of Inference [6]: the if–then rule is paraphrased by an IF relation[8] R from X to Y, i.e. an IFS in $X \times Y$, and B' is computed by taking the $\circ_{\mathcal{T}}$–composition of R and A', defined by, for $y \in Y$

$$B'(y) = R \circ_{\mathcal{T}} A'(y) = \sup_{x\in X} \mathcal{T}(A'(x), R(x,y))) \quad (29)$$

[7] An IFS A in X is called normalized if there exists at least one $x \in X$ such that $A(x) = 1_{L^*}$.

[8] Typically, $R(x,y) = \mathcal{I}(A(x), B(x))$ for some implicator $\mathcal{I}$ on L^* or $R(x,y) = \mathcal{T}(A(x), B(x))$ for some t–norm $\mathcal{T}$ on L^*.

This calculation is computationally costly. In [6], a procedure called inclusion–based approximate reasoning was shown to approximate $B'(y)$ in particular cases.

Theorem 3 *If* $B'(y) = \sup_{x \in X} \mathcal{T}_W(A'(x), \mathcal{I}_{\mathcal{T}_W}(A(x), B(y)))$*, then* $\mathcal{I}_{\mathcal{T}_W}(\mathcal{I}nc_{\mathcal{T}_W}(A', A), B(y)) \geq_{L^*} B'(y)$*, and* $\mathcal{I}nc_{\mathcal{T}_W}(B', B) \geq_{L^*} \mathcal{I}nc_{\mathcal{T}_W}(A', A)$*. Additionally, if the range of* B *is* L^**, then* $\mathcal{I}nc_{\mathcal{T}_W}(B', B) = \mathcal{I}nc_{\mathcal{T}_W}(A', A)$*.*

The approximation is much easier to calculate, since it bypasses the expensive supremum operation. Some promising examples (so far only in fuzzy set theory) have demonstrated the practical use of inclusion–based approximate reasoning. [5]

Note also that the theorem uses $\mathcal{I}nc_{\mathcal{T}_W}$, which satisfies the controversial property (I4b). Another setting in which this measure appears is in the calculation of lower approximations in IF rough set theory as developed in [3].

4.2 Entropy of IFSs

In fuzzy set theory, a common task is to determine the amount of fuzziness in a given fuzzy set (see e.g. [12, 18]). A measure of fuzziness, also called entropy measure, is taken to express the extent to which a crisp distinction between the elements belonging and not belonging to a fuzzy set is lacking, i.e. the extent to which all the elements' membership degrees are close to 0.5. In IFS theory, first steps were taken to define a measure of entropy by Szmidt and Kacprzyk in [16].

To Szmidt and Kacprzyk, an IF entropy measure on X is an $\mathcal{IF}(X) \to [0, 1]$ mapping $\mathcal{E}$ satisfying

1. $\mathcal{E}(A) = 0 \iff A \in \mathcal{P}(X)$
2. $\mathcal{E}(A) = 1 \iff \mu_A = \nu_A$
3. $(\forall x \in X)(\mu_A(x) \leq \mu_B(x) \leq \nu_B(x) \leq \nu_A(x)$ or $\mu_A(x) \geq \mu_B(x) \geq \nu_B(x) \geq \nu_A(x)) \Rightarrow \mathcal{E}(A) \leq \mathcal{E}(B))$
4. $\mathcal{E}(co(A)) = \mathcal{E}(A)$.

These conditions are faithful extensions to those imposed on fuzzy entropy measures[9]. Still, we wish to adjust these requirements in the following respects. First, in a similar vein as for IF inclusion measures, it can be argued that entropy of IFSs cannot be reasonably captured by just one number and is better expressed by elements of L^*. For instance, if $(\forall x \in X)(A(x) = (0, 0))$, then $\mathcal{E}(A)$ should be equal to $(0, 0)$, since no information is available on the fuzziness of A. This also explains our feeling that requirements 2. and 3. are too strong, and should be replaced by $\mathcal{E}(A) = 1_{L^*} \iff (\forall x \in X)(\mu_A(x) = \nu_A(x) = 0.5)$ and $(\forall x \in X)(\mu_A(x) \leq \mu_B(x) \leq 0.5 \leq \nu_B(x) \leq \nu_A(x)$ or $\mu_A(x) \geq \mu_B(x) \geq 0.5 \geq \nu_B(x) \geq \nu_A(x)) \Rightarrow \mathcal{E}(A) \leq_{L^*} \mathcal{E}(B))$.

In fact, we feel that IF entropy should reflect the range of situations that *could* occur if the indeterminacy in A were to disappear, that is: if $\pi_A(x)$ were distributed between $\mu_A(x)$ and $\nu_A(x)$. This idea is illustrated by an example.

[9] Replacing ν_A by $co(\mu_A)$ in the above definition, we can obtain them.

Example 3 *Let A be the IFS in $X = \{x_1, x_2\}$ defined by $A(x_1) = (0.2, 0.3)$, $A(x_2) = (0.1, 0.1)$. Then $\pi_A(x_1) = 0.5$ and $\pi_A(x_2) = 0.8$. The "least fuzzy" fuzzy set obtainable by distributing π_A among μ_A and ν_A is A' defined by $A'(x_1) = (0.2, 0.8), A'(x_2) = (0.9, 0.1)$. The "most fuzzy" fuzzy set derived in this way is A'' defined by $A''(x_1) = (0.5, 0.5), A''(x_2) = (0.5, 0.5)$. Applying an arbitrary fuzzy entropy measure E on X to A' and A'', we have $E(A') \leq E(A'')$. The "real" fuzziness of A is therefore somewhere in the interval $[E(A'), E(A'')]$, and hence could be represented equivalently by $(E(A'), 1 - E(A'')) \in L^*$.*

Kosko [12] was the first to link the entropy of a fuzzy set A to the degree to which $A \cup co(A)$ is included into $A \cap co(A)$, using specific definitions for E and Inc. It is therefore interesting to investigate whether IF entropy measures like the one suggested in the above example can be obtained at all using IF inclusion measures. The following theorem is a very nice affirmation of this conjecture, showing at once the use of frequency–based IF inclusion measures and of t–representable connectives on L^*.

Theorem 4 *Let $\mathcal{E}(A)$, for $A \in \mathcal{IF}(X)$, be defined as*

$$\mathcal{E}(A) = Inc_{\mathcal{S},\mathcal{N}_s}(A \cup co(A), A \cap co(A)) \tag{30}$$

where $\mathcal{S}(x, y) = (\min(1, x_1 + y_1), \max(0, x_2 + y_2 - 1))$. Then $\mathcal{E}$ satisfies the modified conditions of Szmidt and Kacprzyk, and

$$\mathcal{E}(A) = (E(A'), 1 - E(A'')) \tag{31}$$

where A' and A'' are fuzzy sets in X such that for $x \in X$, $A'(x) = \max(1 - \mu_A(x), 1 - \nu_A(x))$, and $A''(x) = \max(0.5, \mu_A(x), \nu_A(x))$, and E is a fuzzy entropy measure, defined for a fuzzy set B in X by

$$E(B) = \frac{2}{|X|} \sum_{x \in X} B \cap co(B)(x) \tag{32}$$

It can be verified that defining e.g. $\mathcal{E}(A) = Inc_{\mathcal{S}_W,\mathcal{N}_s}(A \cup co(A), A \cap co(A))$, a similar decomposition is not possible.

5 Conclusion

This paper has studied various approaches to the definition of intuitionistic fuzzy inclusion measures. We have attempted to reconcile various requirements posed by fuzzy set theory with the specific indeterministic nature of IFSs. This has resulted in two essentially different types of IF inclusion measures, although both are dependent on an implicator on the evaluation set L^*. Future work should focus on the particular meaning of the individual degrees of the result. The example on IF entropy has already shown that under specific conditions on the operations involved a very attractive interpretation can be endowed to the result.

Acknowledgements

The authors would like to thank the anonymous referees for their critical analysis of the paper. Chris Cornelis would like to thank the Fund for Scientific Research–Flanders for funding the research elaborated on in this paper.

References

1. Atanassov, K.T.: Intuitionistic Fuzzy Sets. Physica-Verlag, Heidelberg, New York (1999)
2. Bandler, W., Kohout, L. Fuzzy power sets and fuzzy implication operators. Fuzzy Sets and Systems, **4** (1980) 13–30
3. Cornelis, C., De Cock, M., Kerre, E.: Intuitionistic fuzzy rough sets: on the cross-roads of imperfect knowledge. Accepted to: Expert Systems (2003)
4. Cornelis, C., Deschrijver, G., Kerre, E.E.: Implication in intuitionistic and interval–Valued fuzzy set theory: construction, classification, application. Submitted to: International Journal of Approximate Reasoning (2002)
5. Cornelis, C., Kerre, E.E.: A Fuzzy Inference Methodology Based on the Fuzzification of Set Inclusion. Recent Advances in Intelligent Paradigms and Applications, (A. Abraham, L. Jain, J. Kacprzyk, eds.), Physica-Verlag (2002) 71–89.
6. Cornelis, C., Kerre, E.E.: On the structure and interpretation of an intuitionistic fuzzy expert system. In: Proceedings of EUROFUSE 2002 (B. De Baets, J. Fodor, G. Pasi, eds.) (2002) 173–178.
7. Cornelis, C., Van Der Donck, C., Kerre, E.E.: Sinha–Dougherty approach to the fuzzification of set inclusion revisited. Fuzzy Sets and Systems, **134(2)** (2003) 283–295.
8. Fodor, J., Yager, R.: Fuzzy set theoretic operators and quantifiers. Fundamentals of Fuzzy Sets (D. Dubois, H. Prade, eds.), Kluwer, Boston, Mass. (2000) 125–193
9. Fortemps, P., Slowinski, R.: A graded quadrivalent logic for ordinal preference modelling: Loyola–like approach. Fuzzy Optimization and Decision Making **1** (2002) 93–111
10. Goguen, J.: L–fuzzy Sets. J. Math. Anal. Appl. **18** (1967) 145–174.
11. Kitainik, L.: Fuzzy decision procedures with binary relations. Kluwer, Dordrecht, the Netherlands (1993)
12. Kosko, B.: Fuzzy entropy and conditioning. Kluwer, Dordrecht, the Netherlands (1993)
13. Mendel, J.M.: Uncertain rule–based fuzzy logic systems. Prentice Hall PTR, Upper Saddle River, New Jersey (2001)
14. Sinha, D., Dougherty, E.R.: Fuzzification of set inclusion: theory and applications. Fuzzy Sets and Systems, **55(1)** (1993) 15–42.
15. Stickel, M.E.: Fuzzy four–valued logic for inconsistency and uncertainty. Proceedings of the Eighth International Symposium on Multiple–Valued Logic (1978)
16. Szmidt, E., Kacprzyk, J., Entropy for intuitionistic fuzzy sets. Fuzzy Sets and Systems, **118(3)** (2001), 467–477.
17. Takeuti, G., Titani S.: Intuitionistic fuzzy logic and intuitionistic fuzzy set theory. Journal of Symbolic Logic, **49(3)** (1984) 851-866.
18. Young, V.: Fuzzy subsethood. Fuzzy Sets and Systems, **77(3)** (1996) 371-384.