Fuzzy Inference System (FIS) Extensions Based on Lattice Theory

Vassilis G. Kaburlasos, *Member, IEEE*, and Athanasios Kehagias

***Abstract*—A Fuzzy Inference System (FIS) typically implements a function $f : \mathbb{R}^N \to \mathfrak{T}$, where the domain set $\mathbb{R}$ denotes the totally-ordered set of real numbers, whereas the range set $\mathfrak{T}$ may be either $\mathfrak{T} = \mathbb{R}^M$ (i.e. FIS regressor) or $\mathfrak{T}$ may be a set of labels (i.e. FIS classifier), etc. This work considers the complete lattice $(\mathbb{F}, \preceq)$ of *Type-1 Intervals' Numbers*, or *INs* for short, where an IN F can be interpreted as either a possibility distribution or a probability distribution. In particular, this work concerns the matching degree (or satisfaction degree, or firing degree) part of a FIS. Based on an inclusion measure function $\sigma : \mathbb{F}\times\mathbb{F} \to [0, 1]$ we extend traditional FIS design towards implementing a function $f : \mathbb{F}^N \to \mathfrak{T}$ with the following advantages: (1) accommodation of granular inputs, (2) employment of sparse rules and (3) introduction of tunable (global, rather than solely local) nonlinearities as explained in the manuscript. New theorems establish that an inclusion measure σ is widely (though *implicitly*) used by traditional FISs typically with *trivial* (i.e., *point*) input vectors. A preliminary industrial application demonstrates the advantages of our proposed schemes. Far-reaching extensions of FISs are also discussed.**

***Index Terms* – Fuzzy inference system (FIS), fuzzy lattice reasoning (FLR), granular computing, inclusion measure, fuzzy interval, industrial dispensing, intervals' number (IN), lattice computing (LC)**

I. Introduction

Fuzzy inference systems, or FISs for short, is a long-established technology [24], [53], [72]. A FIS can be interpreted as a fuzzy-logic-based device that implements a function $f : \mathbb{R}^N \to \mathfrak{T}$, where the domain set $\mathbb{R}$ denotes the totally-ordered set of real numbers, whereas the range set $\mathfrak{T}$ may be either $\mathfrak{T} = \mathbb{R}^M$ (i.e. FIS regressor) or $\mathfrak{T}$ may be a set of labels (i.e. FIS classifier), etc. [34]. *Inherent restrictions* of a typical FIS include, first, crisp vector inputs that can not accommodate vagueness, second, a sparse rule base that may not be activated for some system inputs and, third, local (instead of global, as explained below) rule activation that may result in a costly rule proliferation especially when the number of input/output variables increases.

A number of different schemes have been proposed, in various contexts, to overcome aforementioned FIS restrictions. For instance, Zadeh's Compositional Rule of Inference (CRI) [82], applicable to a Mamdani type FIS [53], can deal with fuzzy data. Moreover, granular computing schemes [59] have been proposed for processing non-crisp data as well as for dealing with uncertainty in modeling applications [5], [16]. Type-2 fuzzy sets have been proposed for accommodating vagueness in FISs [54], [77]. Other schemes, such as interpolative reasoning [26], [48], [49] have been proposed for dealing with sparse rule bases. Furthermore, evolving as well as interpretable rule structures have been proposed to encounter rule proliferation [9].

This work concerns the *matching degree* (or satisfaction degree, or firing degree) part of a FIS. In particular, this work proposes a single instrument, namely an *inclusion measure* function $\sigma(.,.)$, towards overcoming all aforementioned "inherent restrictions" of FISs by extending the applicability domain of a typical FIS to the space of *Intervals' Numbers*, or *INs* for short, as explained below. Recall that previous work has employed the term *Fuzzy Interval Number* (*FIN*) instead of the term IN because it stressed a fuzzy interpretation [58]. Moreover, the work in [58] explains that an IN is a mathematical object which may be interpreted as either a probability/possibility distribution or an interval or a real number.

Regarding fuzzy set theory in particular note that even though a fuzzy membership function can be defined on any universe of discourse, it is *fuzzy numbers* (i.e., convex normal fuzzy sets defined on the real numbers $\mathbb{R}$ universe of discourse) which are of special interest due to the widespread use of real numbers [34]. Furthermore, the "resolution identity theorem" [83] has shown that a fuzzy set can, equivalently, be represented either by its membership function or by its α-cuts [47], [57]; obviously, a fuzzy number's α-cut is an interval. This work builds explicitly on the α-cuts representation of fuzzy numbers.

In our previous work we have studied the notion of *generalized intervals* (and *generalized intervals numbers*); these are mathematical objects $[a, b]$ with $a, b \in \mathbb{R}$ where it is not necessary that $a \leq b$. The interested reader can consult [31] and the references included therein. Recently we have turned to "classical intervals" (on which the restriction $a \leq b$ is enforced) and INs[1]. In particular, we have shown that the set $\mathbb{F}$ of INs is a metric lattice [30], [45] with cardinality $\aleph_1$ [33], [34], where "$\aleph_1$" is the cardinality of the set $\mathbb{R}$ of real numbers; moreover, the space $\mathbb{F}$ is a cone in a linear space [36], [58].

In a previous publication, INs have been proposed for extending FISs based on a metric (distance) function [34]. More specifically, a fuzzy membership function was defined in [34] as a function of a metric between INs with the

Vassilis G. Kaburlasos is with the Human-Machines Interaction (HMI) Laboratory, Dept. of Industrial Informatics, TEI of Kavala, 65404 Kavala, Greece (e-mail: vgkabs@teikav.edu.gr).

Athanasios Kehagias is with the Division of Mathematics, Dept. of Math., Phys. & Comp. Sciences, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece (e-mail:kehagiat@auth.gr).

[1] INs, on the one hand, are better suited to certain applications but, on the other hand, require somewhat different methods of analysis.

objective of alleviating the *curse of dimensionality* problem. The same objective can be pursued here by the employment of an *inclusion measure* function. However, only an inclusion measure function extends, in a "principled way" as explained below, the semantics of established FIS practices. In addition, an inclusion measure can extend the applicability of FISs to nonnumeric data domains as discussed below.

Since inclusion measures are central to our approach, let us present some related bibliographic remarks. The literature on inclusion measures is extensive. Hence, we only give a *very brief* introduction (because of space limitations); the interested reader can use our references as a starting point for further study. Fuzzy set inclusion was first defined by Zadeh [81] as a *crisp* relation: a fuzzy set A is either included or not included in another fuzzy set B. Kosko reacted to this by defining set inclusion as a fuzzy relation [50]. This was further enhanced by the *axiomatic* approach; for example Sinha and Dougherty [66] list nine properties that a "reasonable" inclusion measure should have and then derive inclusion measures which have these properties. Other authors [1] obtain inclusion measures from *fuzzy implication operators*. These two approaches (axiomatization and use of fuzzy implications) are combined in several papers [6], [11], [15], [80]; e.g., Burillo et al. [6] introduce a family of implication operators, obtain inclusion measures from these and show that these satisfy Sinha and Dougherty's axioms. A short but very enlightening discussion of the various ways in which "classical subsethood" can be generalized in the fuzzy context appears in [10, pp. 347 and 351-353] where various generalizations of fuzzy subsethood / inclusion measures are categorized into two separate tracks "one logic-based, the other frequency-based". In [7], [46] *lattice-valued* inclusion measures are introduced, i.e., inclusion grades are *partially* ordered. A more common generalization involves real valued inclusion measures which can be applied to *L-fuzzy* sets [18]; specific examples involve *intuitionistic fuzzy sets* [10], [23], [84], [85], *interval-valued* fuzzy sets [84] and *Type-2* fuzzy sets [27], [54]. A quite general class of inclusion measures appropriate for L-fuzzy sets has been recently introduced in [71]. A detailed discussion on the relation between INs and Type-2 fuzzy sets is presented in [42]. The relationship between interval-valued fuzzy sets, intuitionistic fuzzy sets and other extensions of fuzzy sets is discussed in [12], [13]. For some applications of set inclusion see [14], [44], [56] (and the references included therein) as well as the papers discussed in the next paragraph.

In our own early work [29], concentrating on hyperboxes, we have started with a fuzzy measure $\sigma(A, B)$ of the inclusion of a *crisp* set (hyperbox) A into another *crisp* set (hyperbox) B and developed a methodology which uses their inclusion measure for clustering and classification applications [37], [38], [39], [60]. After realizing that the set of hyperboxes in $\mathbb{R}^N$ is lattice-ordered we extended the hyperbox approach to a general lattice data domain as described in the book [31]. In particular, we have used inclusion measures to fuzzify the crisp inclusion relation for (fuzzy) intervals numbers. It turns out that in the lattice of (fuzzy) intervals numbers some technical difficulties arise in the definition of inclusion measures; we address these difficulties in Section III. Let us note in passing that the term "inclusion measure" is probably not general enough; our $\sigma(x, y)$ functions can be better understood as *fuzzy orders*; i.e., $\sigma(x, y)$ expresses the truth value of the statement "$x \leq y$" (where x, y are elements of a lattice). However we stick to the term "inclusion measure" for historical reasons.

The current paper as well as our abovementioned work falls within the general framework of *Lattice Computing* (LC), which has been defined as "the collection of Computational Intelligence tools and techniques that either make use of lattice operators inf and sup for the construction of the computational algorithms or exploit Lattice Theory for language representation and reasoning" [21]. This work adheres to an extended definition of Lattice Computing (LC) that denotes "an evolving collection of tools and mathematical modeling methodologies with the capacity to process lattice ordered data *per se* including logic values, numbers, sets, symbols, graphs, etc" [43], [75]. A recent brief review of selected lattice computing methodologies appears in [20]. Several applications of lattice-theory-based schemes with emphasis on fuzzy control are presented in [28]. An excellent reference on accommodating vagueness and uncertainty in the context of lattice computing is [55]. Specific examples of the lattice computing approach include the connections between granular computing and lattice theory [52], [67] (since information granules are partially/lattice-ordered), lattice-valued (propositional) logics [78], [79], the use of lattice theory to study fuzzy relations [2] and knowledge representations [17] and to extend the notion of a belief function [19]. Also, note that *mathematical morphology* (*MM*), generally conducted in complete lattices or inf-semilattices, is firmly rooted in lattice theory [25], [62], [63]. Hence, morphological neural networks (MNN) including both morphological perceptrons and morphological associative memories (MAMs) [61], [67], [68], [69], [76] should also be classified as lattice computing models. In particular, a fuzzy MAM can be used to implement a FIS based on the complete lattice structure of the class of fuzzy sets [70], [73], [74]. Trends in lattice computing appear in [22], [32], [40].

The paper is organized as follows. Section II presents mathematical preliminaries regarding Intervals' Numbers (INs). Section III details inclusion measure functions with emphasis on INs. Section IV illustrates FIS extensions. Section V presents a preliminary industrial dispensing application. Finally, section VI concludes by summarizing our contribution in perspective. The Appendix includes proofs of Theorems and Lemmas.

II. Mathematical Preliminaries

In this section we present useful definitions, theorems and notation. Since most theorems presented here are "classical", their proofs are omitted.

We use the following set-theoretic notation. The empty set is denoted by $\emptyset$. Both $A \subseteq B$ and $B \supseteq A$ indicate that A is a subset of B; both $A \subset B$ and $B \supset A$ indicate that A is a *proper* subset of B, i.e., there is at least one x such that $x \notin A$ and $x \in B$; both $A \nsubseteq B$ and $B \nsupseteq A$ indicate that A is not a subset of B. Finally, $A \backslash B$ denotes the set of all elements of A which are not contained in B (set difference).

A binary relation $\leq$ on a set P is a *partial order* iff it satisfies three conditions: $x \leq x$ (*reflexivity*), $x \leq y$ and $y \leq x \Rightarrow x = y$ (*antisymmetry*), and $x \leq y$ and $y \leq z \Rightarrow x \leq z$ (*transitivity*). In this case $(P, \leq)$ is called a *partially ordered set* or *poset*. Similarly to the set theoretic notation, $y \geq x$, $x < y$, $y > x$, $x \nleq y, y \ngeq x$ are interpreted in the "obvious" way.

A lattice is a poset $(\mathbb{X}, \leq)$ with the additional property that any two elements $x, y \in \mathbb{X}$ have both an *infimum* (i.e., greatest lower bound) denoted by $x \wedge y$ and a *supremum* (i.e., a least upper bound) denoted by $x \vee y$. It may be the case that for two elements $x, y \in \mathbb{X}$ neither $x \leq y$ nor $x > y$ holds; in this case we say that x and y are *incomparable* and write $x||y$. If in a lattice $(\mathbb{X}, \leq)$ every (x, y) pair satisfies either $x \leq y$ or $x > y$ then we say that lattice $(\mathbb{X}, \leq)$ is *totally ordered*.

Example 2.1: Given any set X, denote by 2^X the set of all subsets of X; then $\left(2^X, \subseteq\right)$ is a (*not* totally ordered) lattice, with set intersection $\cap$ being the infimum operation and set union $\cup$ being the supremum operation.

A lattice $(\mathbb{X}, \leq)$ is called *complete* iff each of its subsets Y has both a greatest lower bound and a least upper bound *in* $\mathbb{X}$ (hence, taking $Y = \mathbb{X}$, we see that a complete lattice has both a *least* element and a *greatest* element).

In this paper we will use a *reference set* $\mathbb{L} \subseteq \overline{\mathbb{R}}$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is the set of *extended real numbers*. We will choose $\mathbb{L}$ so that $(\mathbb{L}, \leq)$ is a complete and totally ordered lattice (here $\leq$ is the "usual" order relation of real numbers). For example, $\mathbb{L}$ can be $\overline{\mathbb{R}}$ itself, or an interval $[a, b] \subset \overline{\mathbb{R}}$, or a finite set $\{x_1, x_2, ..., x_N\} \subset \overline{\mathbb{R}}$. In every case, $\mathbb{L}$ includes a least element denoted by o and a greatest element denoted by i (hence $\mathbb{L} = [o, i]$); the inf and sup operations are denoted by $\wedge$ and $\vee$.

Given $a_1, a_2 \in \mathbb{L}$, with $a_1 \leq a_2$, the *(Type-1) interval* $A = [a_1, a_2]$ is defined by

$$[a_1, a_2] = \{x : x \in \mathbb{L} \text{ and } a_1 \leq x \leq a_2\}.$$

The empty set is also considered an interval, the so-called *empty interval*[2]. We denote the collection of Type-1 intervals of $\mathbb{L}$ (including the empty interval) by $\mathbb{I}(\mathbb{L})$, or simply by $\mathbb{I}$.

The structure $(\mathbb{I}, \subseteq)$ is an ordered set. In fact, it is well-known that the structure $(\mathbb{I}, \subseteq)$ is a complete lattice with respect to the $\subseteq$ order (i.e. set theoretic inclusion). The least element of $\mathbb{I}$ is $\emptyset$, which will also be denoted by O; the greatest element of $\mathbb{I}$ is $\mathbb{L} = [o, i]$, which will also be denoted by I. Given nonempty intervals $A = [a_1, a_2] \in \mathbb{I}$, $B = [b_1, b_2] \in \mathbb{I}$, their infimum and supremum *inside* $\mathbb{I}$ are given by

$$A \cap B = [a_1 \vee b_1, a_2 \wedge b_2] \qquad \text{and} \qquad A \dot{\cup} B = [a_1 \wedge b_1, a_2 \vee b_2].$$

A *fuzzy subset* F of $\mathbb{L}$ is essentially identical to its *membership function* $m_F : \mathbb{L} \to [0, 1]$; intuitively, the number $m_F(x)$ denotes the *degree* to which x belongs to F. A partial order can be defined for fuzzy subsets as follows:

$$F \leq G \Leftrightarrow (\forall x : m_F(x) \leq m_G(x)). \tag{1}$$

(We use, without danger of confusion, the same symbol $\leq$ for the order on real numbers and the one on fuzzy sets). It is easy to check that the infimum (resp. supremum) of two fuzzy sets F, G is a fuzzy set denoted by $F \wedge G$ (resp. $F \vee G$) and defined for every $x \in \mathbb{L}$ by

$$\begin{aligned} m_{F \wedge G}(x) &= m_F(x) \wedge m_G(x), \\ m_{F \vee G}(x) &= m_F(x) \vee m_G(x). \end{aligned} \tag{2}$$

Given a fuzzy subset F with membership function m_F, the h-cut[3] of F is the set

$$F(h) = \{x : m_F(x) \geq h\}.$$

It is well known that a fuzzy subset is fully determined by the family of its h-cuts, i.e., $\{F(h)\}_{h \in [0,1]}$. More specifically, as shown in [57], given a fuzzy set F with membership function m_F, we have:

$$(\forall h : F(h) = G(h)) \Leftrightarrow (\forall x : m_F(x) = m_G(x)).$$

Fuzzy intervals have been studied extensively (for example, see [47], [57] and the references therein). Recall that a *fuzzy interval* is defined as a fuzzy subset F whose every h-cut is an interval: $(\forall h : F(h) \in \mathbb{I})$. We denote the set of all fuzzy intervals by $\mathbb{F}'$.

In [45] it is proved that the set $\mathbb{F}'$ of fuzzy intervals, equipped with the usual fuzzy sets order $\leq$, is a complete lattice; i.e., $(\mathbb{F}', \leq)$ is a complete lattice. The infimum operation is $\wedge$ as defined in (2). The supremum operation is denoted by $\dot{\vee}$ and is defined in terms of membership functions, as follows:

$$m_{F \dot{\vee} G} = \inf\{m_H : H \in \mathbb{F}', F \leq H, G \leq H\}.$$

In words, $F \dot{\vee} G$ is the smallest fuzzy interval which is greater than both F and G.

We now introduce *Type-1 Intervals' Numbers* (*INs* for short).

Definition 2.2: A *Type-1 Intervals' Number* is a function $F : [0, 1] \to \mathbb{I}$ which satisfies

$$\begin{aligned} &F(0) = I, \\ &h_1 \geq h_2 \Rightarrow F(h_1) \subseteq F(h_2), \\ &\forall P \subseteq [0, 1] : \cap_{h \in P} F(h) = F\left(\bigvee P\right). \end{aligned}$$

We denote the class of all (Type-1) INs by $\mathbb{F}$.

Consider the following result, which has been proved in numerous papers and books [3], [47] and also holds in the more general context of L-fuzzy sets [57], [64], [65].

Given an IN $E \in \mathbb{F}$, define $m_{\widetilde{E}}$, the membership function of a fuzzy set $\widetilde{E}$, as follows

$$\forall x : m_{\widetilde{E}}(x) = \sup\{h : x \in E(h)\}.$$

The h-cuts of $m_{\widetilde{E}}$ are denoted by $\widetilde{E}(h)$ and, by definition, satisfy: $\forall h \in [0, 1] : \widetilde{E}(h) = \{x : m_{\widetilde{E}}(x) \geq h\}$. Then, it turns out that for all $h \in [0, 1]$ we have $\widetilde{E}(h) = E(h)$. Hence, $\widetilde{E}$ (the unique fuzzy set with membership function $m_{\widetilde{E}}$) is a *fuzzy interval*. In other words, the h-cuts are the intervals of the original IN and we have a 1-to-1 correspondence between

[2]The empty interval can also be denoted as $[a_1, a_2]$ with any a_1, a_2 such that $a_1 > a_2$.

[3]We use the term "h-cut" instead of the (equivalent) term "α-cut" used in the literature for fuzzy sets. The rationale for introducing the new term stems from two different interpretations for an *Intervals' Number* (IN) as explained in [58].

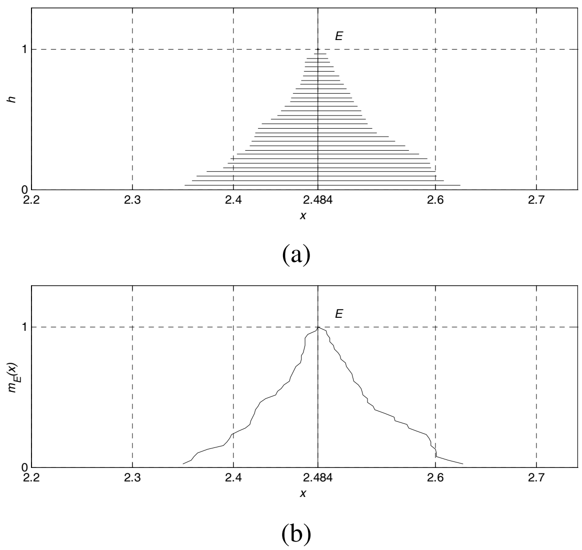


Fig. 1. Two equivalent representations for an Intervals' Number (IN) E include (a) the *interval-representation* $E(h), h \in [0,1]$, and (b) the *membership-function-representation* $m_E(x), x \in \mathbb{R}$. Samples of the former representation are shown here for $L = 32$ different levels spaced evenly over the interval $[0,1]$ on the vertical axis.

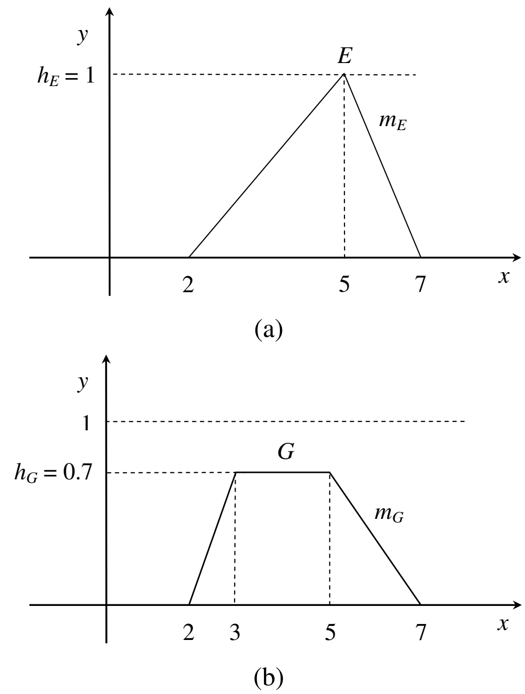


Fig. 2. (a) The height h_E of IN E equals $h_E = 1$. (b) The height h_G of IN G equals $h_G = 0.7$.

fuzzy intervals and INs. There follow two equivalent representations for an IN, namely the *interval-representation* and the *membership-function-representation* (Fig.1). An advantage of the interval-representation is that it enables useful algebraic operations, whereas an advantage of the membership-function-representation is that it enables convenient fuzzy logic interpretations.

The *height* of an IN E, symbolically h_E, is defined as the supremum of the associated membership function $m_E : [-\infty, \infty] \rightarrow [0,1]$ values; i.e., $h_E = \bigvee\limits_{x \in [-\infty,\infty]} m_E(x)$. For example, the height h_E of IN E in Fig.2(a) equals $h_E = 1$, whereas the height h_G of IN G in Fig.2(b) equals $h_G = 0.7$; in particular, note that it is $G(h) = O = \emptyset$ for $h \in (0.7, 1]$.

Just like fuzzy intervals are equipped with a partial order $\leq$, similarly INs can be equipped with a partial order $\preceq$ by defining, for every pair $F, G \in \mathbb{F}$, the relationship $\preceq$ as follows:

$$F \preceq G \Leftrightarrow (\forall h \in [0,1] : F(h) \subseteq G(h)).$$

The isomorphism of $(\mathbb{F}', \leq)$ and $(\mathbb{F}, \preceq)$ is a consequence of the following theorem (the proof of which appears in the Appendix).

Theorem 2.3: For all $F, G \in \mathbb{F}$ we have

$$F \preceq G \Leftrightarrow (\forall h \in [0,1] : F(h) \subseteq G(h)) \Leftrightarrow$$
$$\Leftrightarrow (\forall x \in \mathbb{L} : m_F(x) \leq m_G(x)).$$

Theorem 2.3 has the following corollaries.

Corollary 2.4: For all $F, G \in \mathbb{F}$ the following equivalence holds.

$$F \prec G \Leftrightarrow \left(\begin{array}{c} \forall h : F(h) \subseteq G(h) \\ \exists h_0 : F(h_0) \subset G(h_0) \end{array} \right) \Leftrightarrow$$

$$\Leftrightarrow \left(\begin{array}{c} \forall x : m_F(x) \leq m_G(x) \\ \exists x_0 : m_F(x_0) < m_G(x_0) \end{array} \right).$$

Corollary 2.5: The relationship $\preceq$ is a partial order on $\mathbb{F}$ and $(\mathbb{F}, \preceq)$ is a complete lattice (the *lattice of INs*). If we denote the infimum operation by $\curlywedge$ and the supremum operation by $\dot{\curlyvee}$, then

$$\forall h \in [0,1] : (F \curlywedge G)(h) = F(h) \cap G(h) \text{ and}$$
$$(F \dot{\curlyvee} G)(h) = F(h) \,\dot{\cup}\, G(h).$$

Corollary 2.6: The lattice of fuzzy intervals $(\mathbb{F}', \leq)$ and the lattice of INs $(\mathbb{F}, \preceq)$ are isomorphic.

III. Inclusion Measure Functions

As already mentioned in the Introduction section, an inclusion measure quantifies (by a real number in $[0,1]$) the degree to which a (crisp or fuzzy) set is included in another one. At a higher level of generality, an inclusion measure $\sigma(x, y)$ quantifies the degree to which the order $x \sqsubseteq y$ is true, where x and y are elements of a lattice with order $\sqsubseteq$ (the crisp interval inclusion $\subseteq$ and the fuzzy interval inclusion $\preceq$ are special cases of $\sqsubseteq$). Let us now give a precise definition.

Definition 3.1: Let $(\mathbb{X}, \sqsubseteq)$ be a lattice with inf operation $\sqcap$ and sup operation $\sqcup$. A function $\sigma : \mathbb{X} \times \mathbb{X} \rightarrow [0,1]$ is called an *inclusion measure on* $\mathbb{X}$ if the following properties hold for all $x, y, z \in \mathbb{X}$.

C1 $\sigma(x, x) = 1$.
C2 $x \not\sqsubseteq y \Rightarrow \sigma(x, y) < 1$.
C3 $y \sqsubseteq z \Rightarrow \sigma(x, y) \leq \sigma(x, z)$.

In short, an inclusion measure function $\sigma(x, y)$ quantifies the degree of inclusion of a general lattice element x to another one y, in a "principled way" (in the sense of satisfying properties **C1-C3**). Another way to look at the matter is this: **C1-C3** imply that σ is "compatible" with the order relation $\sqsubseteq$; in fact yet another formulation is that $\sigma(x,y)$ *is* a *fuzzy* order relation. This can also be seen by the following theorem (the proofs of theorems and lemmas are presented in the Appendix).

Theorem 3.2: For all $x, y \in \mathbb{X}$ we have: (a) $x \sqsubseteq y \Leftrightarrow \sigma(x,y) = 1$ and (b) $x \sqcap y \sqsubset x \Leftrightarrow \sigma(x,y) < 1$.

We remark that Definition 3.1 is more general than a previous definition for an inclusion measure [31], [38], [41]; the latter (definition) included the property "**C0** $\sigma(x, O) = 0$, for $x \sqsupset O$" regarding, in particular, a complete lattice $(\mathbb{X}, \sqsubseteq)$ with least element O. However, **C0** is overly restrictive because, for $x \sqsupset O$, it follows $(O \sqsubseteq y$ and $x \sqcap y \sqsubset x) \Rightarrow \sigma(x, O) \leq \sigma(x, y) < 1$; in other words, for $x \sqsupset O$, in a complete lattice, Definition 3.1 only implies $\sigma(x, O) < 1$ instead of the overly restrictive $\sigma(x, O) = 0$.

An inclusion measure function $\sigma : \mathbb{X} \times \mathbb{X} \to [0, 1]$ gives rise to a parametric (fuzzy) membership function $\sigma(., y)$, where y represents a *parameter*. Moreover, recall that an inclusion measure function σ supports two different modes of reasoning, namely *Generalized Modus Ponens* and *Reasoning by Analogy* [41]. Recall also that an employment of inclusion measure function $\sigma(., .)$ for decision-making is called *fuzzy lattice reasoning*, or *FLR* for short [35].

In the rest of this section we will construct inclusion measures specifically for (crisp or fuzzy) intervals. In other words, we will construct inclusion measures on the lattices $(\mathbb{I}, \subseteq)$ and $(\mathbb{F}, \preceq)$. To this end we will use the following two functions, which will be considered fixed for the rest of the section.

A1 A strictly increasing function $v : \mathbb{L} \to [0, \infty)$ which satisfies both $v(o) = 0$ and $v(i) < \infty$.

A2 A strictly decreasing function $\theta : \mathbb{L} \to \mathbb{L}$.

A. Inclusion Measures on $\mathbb{I}$

First we will introduce inclusion measures for crisp intervals. To this end, using functions v and θ, which satisfy **A1-A2**, we introduce *length* functions next.

Definition 3.3: A *length function* $V : \mathbb{I} \to [0, \infty)$ has the following form

$$V(A) = \begin{cases} 0 & \text{iff } A = O \\ v(\theta(a_1)) + v(a_2) & \text{iff } A = [a_1, a_2] \neq O \end{cases}.$$

The following Lemmas describe some properties of length functions, which hold for every v and θ satisfying **A1-A2**.

Lemma 3.4: Every length function V is a strictly increasing function, i.e.,

$$A \subset B \Rightarrow V(A) < V(B)$$

and, for all $A \in \mathbb{I} \setminus \{O\}$, $V(A) > 0$.

Lemma 3.5: For every length function V and for all $A = [a_1, a_2], B = [b_1, b_2], C = [c_1, c_2] \in \mathbb{I}$ we have:

$$O \subset B \subseteq C \Rightarrow V\left(A \dot{\cup} C\right) - V(C) \leq V\left(A \dot{\cup} B\right) - V(B). \quad (3)$$

Now we are ready to introduce functions $\sigma_{\cap}^{V}$ and $\sigma_{\dot{\cup}}^{V}$ which are inclusion measures for every pair (v, θ) which satisfies **A1-A2** (hence we actually define two *families* of inclusion measures, with members of each family determined by the choice of v and θ).

Definition 3.6: The functions $\sigma_{\cap}^{V} : \mathbb{I} \times \mathbb{I} \to [0, 1]$ and $\sigma_{\dot{\cup}}^{V} : \mathbb{I} \times \mathbb{I} \to [0, 1]$ are defined as follows

$$\begin{aligned} \sigma_{\cap}^{V}(A, B) &= \begin{cases} 1 & \text{iff } A = O \\ \frac{V(A \cap B)}{V(A)} & \text{iff } A \neq O \end{cases}, \\ \sigma_{\dot{\cup}}^{V}(A, B) &= \begin{cases} 1 & \text{iff } A \dot{\cup} B = O \\ \frac{V(B)}{V(A \dot{\cup} B)} & \text{iff } A \dot{\cup} B \neq O \end{cases}, \end{aligned} \quad (4)$$

where $V : \mathbb{I} \to [0, \infty)$ is a length function.

Theorem 3.7: The functions $\sigma_{\cap}^{V}$ and $\sigma_{\dot{\cup}}^{V}$ are inclusion measures on $\mathbb{I}$.

Remark 3.8: In previous work [37] we have used an approach similar to the current one to introduce inclusion measures in the lattice of *generalized intervals* (i.e., mathematical objects $[a_1, a_2]$ where we allow $a_1 > a_2$) as follows. Starting with a positive valuation v on the lattice of real numbers, we extended it to a positive valuation V on the lattice of generalized intervals by $V([a_1, a_2]) = v(\theta(a_1)) + v(a_2)$. Then, V can be used to define an inclusion measure on the lattice of generalized intervals. The similarity to our current approach is obvious, but there is a technical difficulty. More specifically, even when v is a valuation (on the real numbers) and θ is a decreasing function, $V([a_1, a_2]) = v(\theta(a_1)) + v(a_2)$ is *not* necessarily a valuation in the lattice of *crisp* intervals. Nevertheless, the significance of Theorem 3.7, is that V can still be used to define inclusion measures, as long as it is a length function.

Example 3.9: In this example we take $\mathbb{L} = [0, M]$, $v(x) = x$ and $\theta(x) = M - x$ (which, obviously, satisfy **A1-A2**). Take intervals $A = [a_1, a_2]$ and $B = [b_1, b_2]$ (in case $A = O = \emptyset$, we write $A = [M, 0]$ and similarly for B). Then

$$V(A) = V([a_1, a_2]) = M + a_2 - a_1.$$

The functions

$$\begin{aligned} \sigma_{\cap}^{V}(A, B) &= \begin{cases} 1 & \text{iff } A = O \\ \frac{V([a_1 \vee b_1, a_2 \wedge b_2])}{V([a_1, a_2])} & \text{otherwise} \end{cases} \quad \text{and} \\ \sigma_{\dot{\cup}}^{V}(A, B) &= \begin{cases} 1 & \text{iff } A = B = O \\ \frac{V([b_1, b_2])}{V([a_1 \wedge b_1, a_2 \vee b_2])} & \text{otherwise} \end{cases} \end{aligned} \quad (5)$$

are inclusion measures on the lattice $(\mathbb{I}([0, M]), \subseteq)$.

Example 3.10: In this example we take $\mathbb{L} = [-\infty, \infty]$ and

$$v(x) = \frac{1}{1 + e^{-\lambda \cdot (x - \mu)}} \text{ and } \theta(x) = 2\mu - x$$

(where $\lambda \in \mathbb{R}^{+}$, $\mu \in \mathbb{R}$) which, obviously, satisfy **A1-A2**. Take intervals $A = [a_1, a_2]$ and $B = [b_1, b_2]$ (in case $A = O = \emptyset$, we write $A = [\infty, -\infty]$ and similarly for B). Then

$$V(A) = \frac{1}{1 + e^{\lambda \cdot (a_1 - \mu)}} + \frac{1}{1 + e^{-\lambda \cdot (a_2 - \mu)}} \quad (6)$$

and the functions of (5) (with V now given by (6)) are inclusion measures on the lattice $(\mathbb{I}([-\infty, \infty]), \subseteq)$.

B. Inclusion Measures on $\mathbb{F}$

We now introduce (families of) inclusion measures for INs.

Definition 3.11: Let $\sigma_{\cap}^{V}$ and $\sigma_{\cup}^{V}$ be the inclusion measure functions of Theorem 3.7 (these definitions depend on the choice of the *length function* V and, ultimately, on both functions v and θ). Now we define the functions *sigma-meet* $\sigma_{\curlywedge}^{V}: \mathbb{F}\times\mathbb{F}\to[0,1]$ and *sigma-join* $\sigma_{\curlyvee}^{V}: \mathbb{F}\times\mathbb{F}\to[0,1]$ as follows.

$$\sigma_{\curlywedge}^{V}(F,G)=\int_0^1 \sigma_{\cap}^{V}(F(h),G(h))dh \qquad \text{and}$$
$$\sigma_{\curlyvee}^{V}(F,G)=\int_0^1 \sigma_{\cup}^{V}(F(h),G(h))dh.$$

Theorem 3.12: The functions $\sigma_{\curlywedge}^{V}$ and $\sigma_{\curlyvee}^{V}$ are inclusion measures on $(\mathbb{F},\preceq)$.

We remark that both inclusion measures $\sigma_{\curlywedge}^{V}$ and $\sigma_{\curlyvee}^{V}$ have been presented elsewhere [30], [31], [35], [42] based on a *positive valuation* function V in the lattice of generalized intervals rather than based on the (different) *length* function V in the lattice $\mathbb{I}$ of intervals as shown in this work.

We argue in the next Section IV that an inclusion measure σ is widely (though *implicitly*) used by traditional FISs. The basis for our claim is provided by the following two theorems.

Theorem 3.13: Take any $F\in\mathbb{F}$ and $X_0\in\mathbb{F}$ such that $X_0(h)=[x_0,x_0]$ for all $h\in[0,1]$. Then, for any length function V, we have $\sigma_{\curlywedge}^{V}(X_0,F)=m_{\widetilde{F}}(x_0)$ (where $\widetilde{F}$ is the fuzzy interval corresponding to IN F).

Remark 3.14: Theorem 3.13 couples an IN's two different representations, namely the interval-representation and the membership-function-representation (see in Fig.1). Note that the proof of Theorem 3.13 justifies our requirement $V(O)=0$ for a *length function* V.

Remark 3.15: Theorem 3.13 can be used to show an interesting connection between inclusion-measure-based inference and the *compositional rule of inference* (CRI) [82]; the latter (CRI) has the form

$$m_G\left(y\right)=\sup_x\left(m_F\left(x\right)\wedge R\left(x,y\right)\right) \tag{7}$$

where m_F, m_G are membership functions and $R\left(x,y\right)$ is a fuzzy relationship connecting x and y. Now suppose $G,F\in\mathbb{F}$; in particular, let F be a trivial IN, i.e., $F(h)=[x_0,x_0]$ for all $h\in[0,1]$ (and a fixed x_0). Furthermore, suppose that for all y, $R\left(x,y\right)$ is a fuzzy interval; the latter corresponds to IN R_y. Now, inferences regarding F can be performed using either the inclusion measure $\sigma_{\curlywedge}^{V}(.,.)$ or CRI. On the one hand, if we use the inclusion measure then by Theorem 3.13 the matching degree is given by

$$\sigma_{\curlywedge}^{V}(F,R_y)=R\left(x_0,y\right).$$

On the other hand, if we use the CRI then by (7) the matching degree is

$$m_G\left(y\right)=\sup_x\left(m_F\left(x\right)\wedge R\left(x,y\right)\right)=R\left(x_0,y\right) \tag{8}$$

since $m_F\left(x\right)=0$ for all $x\neq x_0$ and $m_F\left(x_0\right)=1$. So we see that inclusion-measure-based inference and CRI produce the same result when both F is a trivial IN and $R\left(x,y\right)$ is a fuzzy interval with respect to its first argument. However, for a nontrivial IN F, the CRI and $\sigma_{\curlywedge}^{V}(F,R_y)$ produce different results as demonstrated in the industrial dispensing application example in section V.

Given N lattices $(\mathbb{X}_i,\sqsubseteq_i)$, $i\in\{1,...,N\}$, with the corresponding inf and sup operations denoted by $\sqcap_i$ and $\sqcup_i$, we can define the *product lattice* as follows [4]. The reference set is $\mathbb{X}=\mathbb{X}_1\times\cdots\times\mathbb{X}_N$; for any N-tuples $\mathbf{x}=(x_1,\ldots,x_N)\in\mathbb{X}$ and $\mathbf{y}=(y_1,\ldots,y_N)\in\mathbb{X}$, the order $\sqsubseteq$ is defined by: $\mathbf{x}\sqsubseteq\mathbf{y}\Leftrightarrow(\forall i\in\{1,...,N\}: x_i\sqsubseteq_i y_i)$. Then $(\mathbb{X},\sqsubseteq)$ is a lattice with inf $\sqcap$ and sup $\sqcup$ operations defined as follows:

$$\mathbf{x}\sqcap\mathbf{y}=(x_1\sqcap_1 y_1,...,x_N\sqcap_N y_N) \text{ and}$$
$$\mathbf{x}\sqcup\mathbf{y}=(x_1\sqcup_1 y_1,...,x_N\sqcup_N y_N).$$

The following definition and theorem show how to introduce inclusion measures to "product" or, equivalently, "aggregate" lattices.

Definition 3.16: Let lattice $(\mathbb{X},\sqsubseteq)$ be the product of N lattices $(\mathbb{X}_i,\sqsubseteq_i)$ $(i\in\{1,...,N\})$ and suppose σ_i is an inclusion measure on $(\mathbb{X}_i,\sqsubseteq_i)$ (for $i\in\{1,2,...,N\}$). We define functions $\sigma_\wedge:\mathbb{X}\times\mathbb{X}\to[0,1]$ and $\sigma_\Pi:\mathbb{X}\times\mathbb{X}\to[0,1]$ as follows

$$\sigma_\wedge(\mathbf{x},\mathbf{y})=\min_{i\in\{1,...,N\}}\sigma_i(x_i,y_i) \text{ and}$$
$$\sigma_\Pi(\mathbf{x},\mathbf{y})=\prod_{i=1}^{N}\sigma_i(x_i,y_i).$$

Theorem 3.17: The functions $\sigma_\wedge(\mathbf{x},\mathbf{y})$ and $\sigma_\Pi(\mathbf{x},\mathbf{y})$ are inclusion measures on the product lattice $(\mathbb{X},\sqsubseteq)$.

Remark 3.18: Any one of the lattices $(\mathbb{X}_i,\sqsubseteq_i)$ implicit in Theorem 3.17 can be a lattice of crisp intervals or INs (or, in fact, any other lattice) and the inclusion measures σ_i can be any of the previously defined $\sigma_{\cap}^{V}$, $\sigma_{\cup}^{V}$, $\sigma_{\curlywedge}^{V}$, $\sigma_{\curlyvee}^{V}$ (for various choices of functions v_i,θ_i). We can use these "component lattices" to build an "aggregate lattice"; then Theorem 3.17 tells us how to obtain an inclusion measure for this aggregate lattice. Furthermore, we point out that Theorem 3.13 applies, in particular, to the lattice $(\mathbb{F},\preceq)$ of INs, whereas Theorem 3.17 applies to a general product lattice $(\mathbb{X},\sqsubseteq)$.

C. Some Remarks on the Construction of the Inclusion Measures

We now present some remarks about our methodology of constructing inclusion measures in $(\mathbb{F},\preceq)$. This methodology consists of two steps: in the first step we construct $\sigma_{\cap}^{V}$ and $\sigma_{\cup}^{V}$, inclusion measures for *crisp* intervals; in the second step we construct $\sigma_{\curlywedge}^{V}$ and $\sigma_{\curlyvee}^{V}$, inclusion measures for Intervals' Numbers.

Regarding $\sigma_{\cap}^{V}$ and $\sigma_{\cup}^{V}$, note that these are determined by the length function $V\left([a_1,a_2]\right)=v\left(\theta\left(a_1\right)\right)+v\left(a_2\right)$ and so ultimately by the strictly increasing function $v:\mathbb{L}\to[0,\infty)$ and the strictly decreasing function $\theta:\mathbb{L}\to\mathbb{L}$. For instance, as shown in Example 3.9, given $v(x)=x$ as well as $\theta(x)=M-x$ it follows $V([a,b])=M+b-a$. In practice, a *parametric* family of functions $v(.)$ and/or $\theta(.)$ is proposed by the user. Note that different authors have already proposed linear /hyperbolic tangent /arctan /sigmoid positive valuation functions $v(.)$ [36], [37], [41], [51], [52]; whereas, the corresponding function $\theta(.)$ is an affine. Then, typically, optimal parameter estimates are induced from "training data" using stochastic search techniques (e.g., genetic algorithms) as demonstrated in numerous pattern classification and regression applications [36], [42], [58].

Function $V([a_1, a_2])$ is meant (as its name indicates) as a generalization of the *length* of the interval $[a_1, a_2]$. In this light, for example, the inclusion of interval $A = [a_1, a_2]$ in interval $B = [b_1, b_2]$ is a ratio of lengths ($\sigma_{\cap}^{V}(A, B) = \frac{V(A \cap B)}{V(A)}$), namely the length of the common part of A and B divided by the total length of A. This approach to inclusion measure, not just for crisp intervals but for *general fuzzy sets*, has been introduced in [50]; see [80] for many interesting generalizations. We can also understand $V(.)$ as a *probability measure*, in which case $\frac{V(A \cap B)}{V(A)}$ will be understood as a conditional probability. Now, recall that a probability is a special case of a *set-measure*[4] which, in turn, is a generalization of length. Similar (though not identical) remarks can be made about $\sigma_{\cup}^{V}(A, B) = \frac{V(B)}{V(A \cup B)}$. In short: *we obtain our inclusion measures by generalizing the concept of length.*

It remains to explain why $V([a_1, a_2]) = v(\theta(a_1)) + v(a_2)$ is indeed a generalization of length. Recall that we work with crisp *intervals*; these are a restricted (but very useful) type of sets which are characterized by two numbers: their endpoints. Hence $V([a_1, a_2])$ need only depend on the endpoints a_1, a_2; and if it is meant to generalize length, then $V([a_1, a_2])$ must be increasing with a_2 and decreasing with a_1; an easy way to achieve this is by setting $V([a_1, a_2]) = v(\theta(a_1)) + v(a_2)$.

We now turn to $\sigma_{\curlywedge}^{V}$ and $\sigma_{\curlyvee}^{V}$. These are inclusion measures for INs (or, equivalently, for *fuzzy* intervals) and they work by *aggregating* the degrees of inclusion for an infinite family of crisp intervals, namely the cuts $A(h)$ and $B(h)$ for every h value. A natural way to achieve this aggregation is by using the integral operator; this is the motivation behind Definition 3.11.

Let us conclude by remarking that the length function can be generalized in other ways. Perhaps the simplest one is to let $V([a_1, a_2])$ be a true set-measure. For example, one could try to obtain a family of inclusion measures by using

$$V([a_1, a_2]) = \int_{a_1}^{a_2} w(x)\, dx \tag{9}$$

where $w(x)$ is a strictly positive bounded function [33], [34]. However (9) yields $V([a_1, a_1]) = 0$, which contradicts the requirement $V(A) = 0 \Rightarrow A = O$. More generally, an inclusion measure cannot be obtained from a set-measure V under which exist nonempty sets of measure zero. This technical difficulty can be resolved on discrete spaces, in which set-measures can be used to construct inclusion measures. For example let $\mathbb{L} = \{x_1, ..., x_N\}$ be a subset of the real numbers (equipped with the "usual" order $\leq$) and define (for $n = 1, ..., N$) the "weights" $v(x_n) = w_n > 0$; then the function

$$V([a_1, a_2]) = \sum_{a_1 \leq x_n \leq a_2} v(x_n)$$

can be used to construct an inclusion measure in a manner similar to that of Section III-A. Further generalizations are possible; we will pursue this direction in a future publication.

[4]As in "Lebesgure measure".

IV. Fuzzy Inference System (FIS) Extensions

Even though an explicit connection was shown between *mathematical lattices* and *fuzzy sets* since the introduction of fuzzy set theory [81], it is remarkable, as explained in [34], that no tools have been established for FIS analysis and design based on lattice theory. In this connection we have presented two theoretical contributions, that is, Theorem 3.13 and Theorem 3.17, which substantiate that inclusion measures σ are widely (though *implicitly*) used by traditional FISs as detailed in this section.

Here is an interesting consequence of Theorem 3.13. Take a fuzzy interval $F \in \mathbb{F}$ and its corresponding membership function $m_F : [-\infty, \infty] \to [0, 1]$. Then m_F may, equivalently, be represented by the inclusion measure function $\sigma_{\curlywedge}^{V}(X, F) = m_F(x)$ for trivial INs $X = X(h) = [x, x]$, where $h \in [0, 1]$ and $x \in [-\infty, \infty]$. Parameter "F" of the fuzzy set $\sigma_{\curlywedge}^{V}(X, F)$ is called here *kernel* (of the fuzzy set $\sigma_{\curlywedge}^{V}(X, F)$). In terms of fuzzy set theory, the kernel F constitutes the *core* of the fuzzy set $\sigma_{\curlywedge}^{V}(X, F)$.

An inclusion measure $\sigma(X, F)$ has a significant potential in FIS applications due to several advantages. First, for any membership function $m_F : [-\infty, \infty] \to [0, 1]$ both inclusion measures $\sigma_{\curlywedge}^{V}(X, F)$ and $\sigma_{\curlyvee}^{V}(X, F)$ can accommodate vagueness in X in a "principled way", in the sense of satisfying the properties **C1-C3** of Definition 3.1. A second advantage, in particular for inclusion measure $\sigma_{\curlyvee}^{V}(X, F)$, is its applicability beyond the support of the fuzzy set F. A third advantage for both inclusion measures $\sigma_{\curlywedge}^{V}(X, F)$ and $\sigma_{\curlyvee}^{V}(X, F)$ is their (parametric) tunability since both $\sigma_{\curlywedge}^{V}(X, F)$ and $\sigma_{\curlyvee}^{V}(X, F)$ are defined based on parametric functions $\theta : \mathbb{L} \to \mathbb{L}$ and $v : \mathbb{L} \to [0, \infty)$.

Furthermore, it is known that a traditional FIS typically uses either the "min" operator or the "product" operator to calculate the degree of truth of a fuzzy rule (involving N simple propositions as antecedents) from the degrees of truth of the aforementioned N propositions. Theorem 3.17 establishes that a traditional FIS *implicitly* employs inclusion measure functions $\sigma_{\wedge}(\mathbf{x}, \mathbf{y})$ and $\sigma_{\Pi}(\mathbf{x}, \mathbf{y})$ for the "min" operator and the "product" operator, respectively. Therefore, an *explicit* employment of an inclusion measure is expected to result in the three abovementioned advantages as demonstrated below. We point out explicitly that this work is not concerned with the consequents of rules; instead, our interest here focuses on rule antecedents as explained in the following examples, where all the definite integrals were calculated by numerical integration using a standard commercial software package (MATLAB).

Example 4.1: Fig. 3 displays the antecedent of a typical FIS rule, say R. In the interest of simplicity, without loss of generality, we show only two INs E_1 and E_2 with parabolic membership functions $m_{E_1}(x) = -x^2 + 6x - 8$ and $m_{E_2}(x) = -0.25x^2 + 3.5x - 11.25$, respectively. Let an input $(x_{1,0}, x_{2,0}) = (3.5, 5.5)$ be presented to the rule R as shown in Fig. 4(a). Using traditional FIS techniques, the activation $m_R(x_{1,0}, x_{2,0})$ of rule R is a function of both numbers $m_{E_1}(x_{1,0}) = 0.75$ and $m_{E_2}(x_{2,0}) = 0.4375$; the latter (numbers) are the degrees of membership of the inputs $x_{1,0}$ and $x_{2,0}$ to the INs E_1 and E_2, respectively. Popular functions $m_R(.,.)$ in the literature include $m_{R1}(x_{1,0}, x_{2,0}) = \mathtt{min}\{m_{E_1}(x_{1,0}), m_{E_2}(x_{2,0})\}$ and $m_{R2}(x_{1,0}, x_{2,0}) = m_{E_1}(x_{1,0}) \cdot m_{E_2}(x_{2,0})$. Recall that the advantage of the former function $m_{R1}(.,.)$ is that it is computed quickly, whereas the advantage of the latter function $m_{R2}(.,.)$ is that it results in a "smooth" output (without abrupt changes). Identical results were obtained using inclusion measure $\sigma_{\curlywedge}^{V}(.,.)$ with $\sigma_{\wedge}(\mathbf{X_0}, \mathbf{E})$ and $\sigma_{\Pi}(\mathbf{X_0}, \mathbf{E})$, respectively, where $\mathbf{E} = (E_1, E_2)$ and $\mathbf{X_0} = (X_{1,0}, X_{2,0})$ with $X_{1,0} = X_{1,0}(h) = (x_{1,0}, x_{1,0}) = (3.5, 3.5)$ and $X_{2,0} = X_{2,0}(h) = (x_{2,0}, x_{2,0}) = (5.5, 5.5)$, for all $h \in [0, 1]$. In conclusion, the results by $\sigma_{\curlywedge}^{V}(.,.)$ do not differ from the results by traditional FIS techniques. In addition, our proposed technology can overcome the abovementioned *inherent* FIS restrictions as follows.

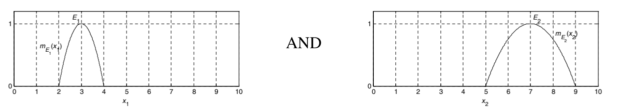


Fig. 3. The antecedent of a typical FIS rule R including the conjunction of two propositions, namely "*variable* V_1 is E_1" and "*variable* V_2 is E_2". The membership functions of INs E_1 and E_2 are the parabolas $m_{E_1}(x_1)$ and $m_{E_2}(x_2)$, respectively.

Example 4.2: An inclusion measure can accommodate granular input INs towards representing vagueness in practice. For instance, consider the granular input INs X_1 and X_2 shown in Fig. 4(b) each with an isosceles (triangular) membership function of width $2 \cdot 0.2 = 0.4$ centered at $x_1 = 3.5$ and $x_2 = 5.5$, respectively. Given functions $v(x) = x$ and $\theta = 10 - x$ over the domain $[0, 10]$, it follows $\sigma_{\curlywedge}^{V}(X_1, E_1) = \int_0^{0.6825} 1dh + \int_{0.6825}^{0.7902} \frac{48.5-h+5\sqrt{1-h}}{52-2h}dh + \int_{0.7902}^{1} 0dh \approx 0.7898$ and $\sigma_{\curlywedge}^{V}(X_2, E_2) = \int_0^{0.3331} 1dh + \int_{0.3331}^{0.5088} \frac{43.5-h+10\sqrt{1-h}}{52-2h}dh + \int_{0.5088}^{1} 0dh \approx 0.5072$. Note that the upper integral ends 0.7902 and 0.5088 are upper bounds for Zadeh's Compositional Rule of Inference (CRI) [82], only for a Mamdani type FIS as explained next. For a traditional FIS rule "if A then B", symbolically $A \Rightarrow B$, represented by a fuzzy relation R, the (fuzzy) output B' to a fuzzy input A' can be computed using Zadeh's CRI: $B' = A' \circ (A \Rightarrow B) = A' \circ R$, where the *max-min* product "A' and R" in [82] was later generalized by the "sup T" compositional operator. On the one hand, restrictions of Zadeh's CRI include, first, the aforementioned fuzzy sets A and A' need to overlap, otherwise a zero fuzzy output B' results in; second, a fuzzy relation R can be defined for a Mamdani type FIS [53] but not for a Sugeno type FIS [72] – Recall that a Mamdani type FIS has been described as a function $m : \mathbb{F}^N \to \mathbb{F}^M$, whereas a Sugeno type FIS has been described as a function $s : \mathbb{F}^N \to \mathcal{P}_p$, where $\mathcal{P}_p$ is a family of models with p parameters [34]. On the other hand, since an inclusion measure involves only rule antecedents, an inclusion measure is applicable on either Mamdani- or Sugeno- type FISs. In particular, inclusion measure $\sigma_{\curlyvee}(.,.)$ may involve nonoverlapping INs as demonstrated next.

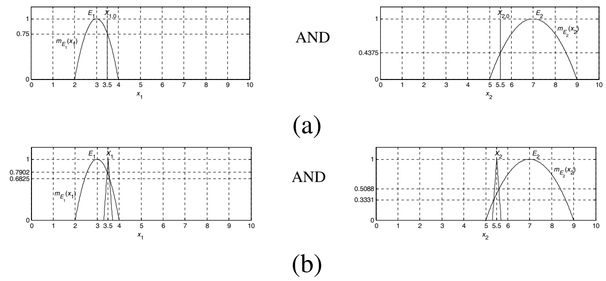


Fig. 4. Consider the antecedent of rule R from Fig.3. (a) Rule R is activated by a trivial INs vector $\mathbf{X_0} = (X_{1,0}, X_{2,0})$. Using either a traditional FIS or inclusion measure $\sigma_{\curlywedge}(.)$ the degree of truth of proposition "*variable* $V_1 = X_{1,0}$ is E_1" equals 0.75; furthermore, the degree of truth of proposition "*variable* $V_2 = X_{2,0}$ is E_2" equals 0.4375. (b) Rule R is activated by a nontrivial INs vector $\mathbf{X} = (X_1, X_2)$ such that each one of INs X_1 and X_2 has an isosceles (triangular) membership function of width $2 \cdot 0.2 = 0.4$. Only inclusion measure $\sigma_{\curlywedge}(.)$ can calculate, as explained in the text, the degrees of truth of the propositions "*variable* $V_1 = X_1$ is E_1" and "*variable* $V_2 = X_2$ is E_2"; in particular, it was computed $\sigma_{\curlywedge}(X_1 \preceq E_1) \approx 0.7898$ and $\sigma_{\curlywedge}(X_2 \preceq E_2) \approx 0.5072$, respectively.

Example 4.3: Fig. 5(a) shows a trivial INs input vector $\mathbf{X_0} = (X_0, X_0)$ beyond rule support, where $X_0 = X_0(h) = (4.5, 4.5)$, $h \in [0, 1]$. Given functions $v(x) = x$ and $\theta = 10 - x$, it follows that $\sigma_{\curlyvee}^{V}(X_0, E_1) = \int_0^1 \frac{10+2\sqrt{1-h}}{11.5+\sqrt{1-h}}dh \approx 0.9311$ and $\sigma_{\curlyvee}^{V}(X_0, E_2) = \int_0^1 \frac{10+4\sqrt{1-h}}{12.5+2\sqrt{1-h}}dh \approx 0.9144$. Fig. 5(b) shows a nontrivial INs input vector $\mathbf{X} = (X, X)$, also beyond rule support, where IN X has an isosceles (triangular) membership function of width $2 \cdot 0.2 = 0.4$ centered at 4.5. It follows that $\sigma_{\curlyvee}^{V}(X, E_1) = \int_0^1 \frac{50+10\sqrt{1-h}}{58.5-h+5\sqrt{1-h}}dh \approx 0.9235$, and $\sigma_{\curlyvee}^{V}(X, E_2) = \int_0^1 \frac{50+20\sqrt{1-h}}{63.5-h+10\sqrt{1-h}}dh \approx 0.9078$. We remark that computing a rule activation beyond rule support is important for decision-making in a *sparse* rule base. Next, we discuss how traditional FISs typically handle a sparse rule base. Inference in sparse rule bases is typically carried out by Fuzzy Rule Interpolation (FRI) [48], [49] motivated towards reducing a fuzzy model's rule complexity by inducing fuzzy rules from other ones according to the following scheme [8]:

Rule 1:
IF (X_1 is A_{11}) and ... and (X_m is A_{1m}) THEN Y is B_1
⋮
Rule n:
IF (X_1 is A_{n1}) and ... and (X_m is A_{nm}) THEN Y is B_n
Observation:
(X_1 is A_1^*) and ... and (X_m is A_m^*)

Conclusion: Y is B^*,

where X_j is an *antecedent* variable (or, equivalently, *system input* variable), Y is the *consequent* variable (or, equivalently, *system output* variable), A_{ij} is a fuzzy number value for variable X_j and B_i is a fuzzy number value for variable Y, $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$.

A number of FRI schemes have been introduced in the literature [8], [26], [48], [49]. In general, the FRI techniques proposed in the literature are restricted to Mamdani type FISs, where nonlinearities are introduced by *ad hoc* function $f : \mathbb{F}^N \rightarrow \mathbb{F}$ FIS techniques (for details the reader may refer to [34]) without, usually, a capacity for "fine tuning". Even though (non)linear rule interpolation/extrapolation is feasible in the cone $\mathbb{F}$ of INs [36], [58], this work deals with sparse rules differently, with significant advantages. More specifically, we treat a fuzzy number A (with an arbitrary membership function shape) in a sparse rule base, as the kernel of the fuzzy set $\sigma^V_{\curlyvee}(X, A)$, where V is a length function with tunable parameters. That is, instead of inserting new fuzzy rules by interpolation/extrapolation, we extend the support of the existing (sparse) rules. Since this work focuses on the *matching degree* (or satisfaction degree, or firing degree) part of a FIS, our techniques here unify the treatment of Mamdani type FISs and Sugeno type FISs. The number of the "closest", in an inclusion measure sense, rules to "fire" is user-defined. How exactly to use an inclusion measure, e.g. towards computing multiple "firing rules" and/or resolving any inconsistencies, depends on a specific application and it is a topic for future work.

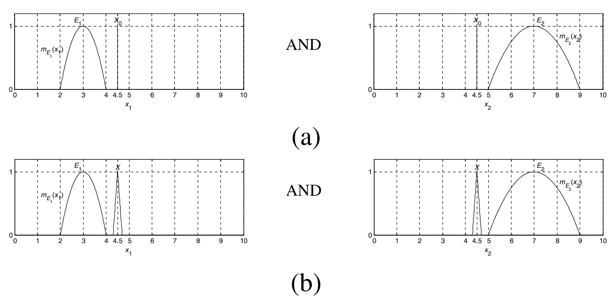


Fig. 5. Consider the antecedent of rule R from Fig.3. (a) A trivial INs input vector $\mathbf{X_0} = (X_0, X_0)$ is presented. (b) A nontrivial INs input vector $\mathbf{X} = (X, X)$ is presented such that IN X has an isosceles (triangular) membership function of width $2{\cdot}0.2 = 0.4$. Neither a traditional FIS nor inclusion measure $\sigma_\curlywedge(.)$ can activate rule R because input IN X (as well as input IN X_0) is outside the support of both IN E_1 and IN E_2. Nevertheless, inclusion measure $\sigma_{\dot\curlyvee}(.)$ can activate rule R. In particular, it was computed $\sigma_{\dot\curlyvee}(X_0 \preceq E_1) \approx 0.9311$ and $\sigma_{\dot\curlyvee}(X_0 \preceq E_2) \approx 0.9144$; moreover, $\sigma_{\dot\curlyvee}(X \preceq E_1) \approx 0.9235$ and $\sigma_{\dot\curlyvee}(X \preceq E_2) \approx 0.9078$.

Example 4.4: An inclusion measure can employ alternative functions than functions $v(x) = x$ and $\theta(x) = 10 - x$ employed above. More specifically, in Fig.6, we considered the INs E_1, E_2, X_0 and X from Fig.5. Moreover, we considered both the (sigmoid) strictly increasing function $v_s(x; \lambda, \mu) = \frac{1}{1+e^{-\lambda(x-\mu)}}$, $x \in \mathbb{R}$, where $\lambda \in \mathbb{R}^+, \mu \in \mathbb{R}$, and the strictly decreasing function $\theta(x; \mu) = 2\mu - x$. Several inclusion measure values were computed and the corresponding results are displayed in Table I for various values of λ and μ. Next, we computed all aforementioned inclusion measure values using the same (sigmoid) strictly increasing function $v_s(x; \lambda, \mu)$, nevertheless we used the strictly decreasing function $\theta(x) = -x$ instead; the corresponding results are displayed in Table II. Tables I and II demonstrate that different functions $v(.)$ and $\theta(.)$ may result in different fuzzy sets $\sigma_{\dot\curlyvee}(X, M)$ with the same kernel M. Most interesting is that inequality $\sigma^V_{\curlyvee}(X_0, E_1) > \sigma^V_{\curlyvee}(X_0, E_2)$ in Table I is reversed in Table II. That is, Tables I and II demonstrate that (parametric) functions $v(.)$ and $\theta(.)$ can be used as instruments for tunable decision-making. Note that conventional FISs carry out solely *local* rule activation in the sense that a rule is activated if and only if an input falls inside its (rule) support, whereas a FIS based on inclusion measure $\sigma^V_{\curlyvee}(., .)$ can carry out *global* rule activation in the sense that a rule can be activated for any input either inside or outside its (rule) support. In conclusion, conventional FISs can introduce only local nonlinearities typically by tuning the shape and/or the location of fuzzy sets involved in the computations, whereas the proposed FISs, based on an inclusion measure, can, in addition, introduce global nonlinearities via the tunable (parametric) functions $v(.)$ and $\theta(.)$.

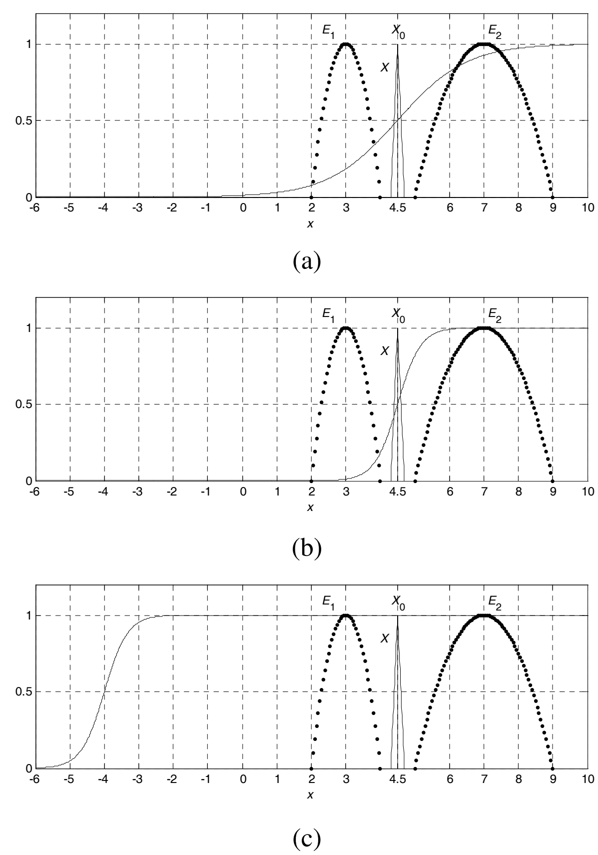


Fig. 6. Parabolic INs E_1 and E_2 (in dotted lines) are displayed as well as both trivial IN X_0 and triangular IN X from Fig.5. Inclusion measure $\sigma_{\dot\curlyvee}(.)$ values were computed using the displayed sigmoid strictly increasing functions $v_s(x; \lambda, \mu) = 1/(1 + e^{-\lambda(x-\mu)})$ for different values of the parameters λ and μ including: (a) $\lambda = 1$, $\mu = 4.5$; (b) $\lambda = 3$, $\mu = 4.5$; (c) $\lambda = 3$, $\mu = -4$. The corresponding $\sigma_{\dot\curlyvee}(.)$ values for the strictly decreasing functions $\theta(x) = 2\mu - x$ and $\theta(x) = -x$ are displayed in Table I and Table II, respectively.

V. An Industrial Dispensing Application

This section demonstrates an employment of our proposed techniques in a preliminary industrial application regarding liquid dispensing. The industrial problem as well as a software application platform, namely XtraSP.v1, and algorithm CALCIN have been detailed elsewhere [35].

TABLE I
INCLUSION MEASURE VALUES REGARDING FIG.6. THE STRICTLY DECREASING FUNCTION $\theta(x) = 2\mu - x$ WAS EMPLOYED.

Inclusion Measure	Fig.6(a): $v_s(x; 1, 4.5)$	Fig.6(b): $v_s(x; 3, 4.5)$	Fig.6(c): $v_s(x; 3, -4)$
$\sigma_{\curlyvee}(X_0, E_1)$	0.8598	0.7261	1
$\sigma_{\curlyvee}(X_0, E_2)$	0.8287	0.7036	1
$\sigma_{\curlyvee}(X, E_1)$	0.8446	0.6916	1
$\sigma_{\curlyvee}(X, E_2)$	0.9449	0.9148	1

TABLE II
INCLUSION MEASURE VALUES REGARDING FIG.6. THE STRICTLY DECREASING FUNCTION $\theta(x) = -x$ WAS EMPLOYED.

Inclusion Measure	Fig.6(a): $v_s(x; 1, 4.5)$	Fig.6(b): $v_s(x; 3, 4.5)$	Fig.6(c): $v_s(x; 3, -4)$
$\sigma_{\curlyvee}(X_0, E_1)$	0.6114	0.1793	1
$\sigma_{\curlyvee}(X_0, E_2)$	0.9999	1	0.8571
$\sigma_{\curlyvee}(X, E_1)$	0.5803	0.1509	1
$\sigma_{\curlyvee}(X, E_2)$	1	1	0.9626

A. *Feedback Control Based on Fuzzy Lattice Reasoning*

Effective control in the food industry calls for sensible decision-making rather than for ultimate precision. Therefore, we estimated the volume of a liquid being dispensed to a mixing tank by both flowmeter measurements and ultrasonic level meter (U.L.M.) measurements accommodating vagueness as follows.

Even though the flowmeter device supplies one precise measurement, there is uncertainty regarding the dispensed volume due to both time-delays and the (exact) storage capacity of the pipes/devices used to drive a liquid to the mixing tank. The latter uncertainty has been modeled by two adjacent uniform pdfs, respectively, one above- and the other below- a flowmeter measurement [35]. Hence, in our computer simulation experiments below, five numbers were drawn randomly (uniformly) for each one of the aforementioned two pdfs. In addition, in a short sequence, we considered randomly (uniformly) ten successive measurements of the liquid level in the mixing tank using the U.L.M. device. In conclusion, we kept inducing an IN V from a population of twenty measurements using algorithm CALCIN [35], [42]. In our experiments, for any population of twenty measurements, we assumed an average measurements population range of 6 $[lt]$ with an insignificant standard deviation.

The following simple decision-making rule was assumed for dispensing a liquid to the mixing tank.

Rule R_0 : IF the volume V (of the liquid being dispensed) is V_{ref} THEN *stop dispensing*

We remark that $V_{ref}(h) = [V_0 - \Delta V, V_0 + \Delta V], h \in [0, 1]$, where "$V_0$" is the desired (crisp) volume and "ΔV" is an acceptable tolerance regarding the desired liquid volume V_0.

Fig. 7 shows the feedback control scheme we employed towards automating industrial liquid dispensing. We assumed that the degree of fulfilment of rule R_0 equals the degree of truth of its antecedent. The latter degree of truth equals the degree of membership of IN V to the fuzzy inclusion measure function $\sigma(V, V_{ref})$ with kernel $V_{ref} = V_{ref}(h) = [V_0 - \Delta V, V_0 + \Delta V], h \in [0, 1]$. We *stop dispensing* when the degree of truth of the antecedent statement "the volume V (of the liquid being dispensed) is V_{ref}" is larger-than or equal-to a user-defined threshold $T \in [0, 1]$. An advantage of the proposed scheme is its capacity to deal in a "principled way", in the sense of satisfying properties **C1-C3** of Definition 3.1, with vagueness in both the system output and reference signals represented by INs V and V_{ref}, respectively.

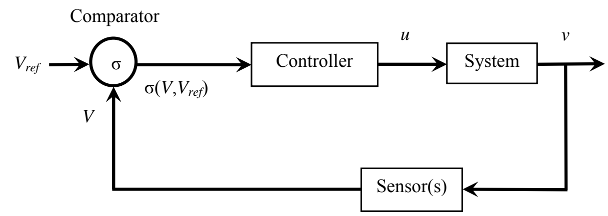


Fig. 7. Feedback control based on fuzzy lattice reasoning (FLR): The system output (volume) $v \in \mathbb{R}$ is sampled by sensor(s); the produced population of measurements is represented by IN V. The latter (V) is fed back for comparison to the reference volume IN $V_{ref} = V_{ref}(h) = [1496.4, 1503.6], h \in [0, 1]$. An inclusion measure $\sigma(V, V_{ref})$ drives the controller who generates a binary (ON/OFF) control signal $u \equiv \{\sigma(V, V_{ref}) \geq T\}$, where $T \in [0, 1]$ is a user-defined threshold.

B. *Application of our Techniques*

Inclusion measure *sigma-meet* $\sigma_{\curlywedge}(V, V_{ref})$ was computed as follows:

$$\sigma^V_{\curlywedge}(V, V_{ref}) = \int_0^1 \sigma^V_{\cap}([a_h, b_h], [V_0 - \Delta V, V_0 + \Delta V])dh = \int_0^{m_V(V_0-\Delta V) \vee m_V(V_0+\Delta V)} \frac{v_s(\theta(a_h \vee (V_0-\Delta V))) + v_s(b_h \wedge (V_0+\Delta V))}{v_s(\theta(a_h)) + v_s(b_h)} dh,$$

where the symbol $m_V(.)$ above denotes the membership function of IN $V = V(h) = [a_h, b_h]$, $h \in [0, 1]$.

Inclusion measure *sigma-join* $\sigma_{\curlyvee}(V, V_{ref})$ was computed as follows:

$$\sigma^V_{\curlyvee}(V, V_{ref}) = \int_0^1 \sigma^V_{\cup}([a_h, b_h], [V_0 - \Delta V, V_0 + \Delta V])dh =$$
$$= \int_0^1 \frac{v_s(\theta(V_0-\Delta V)) + v_s(V_0+\Delta V)}{v_s(\theta(a_h \wedge (V_0-\Delta V))) + v_s(b_h \vee (V_0+\Delta V))} dh.$$

We sought an optimal estimation of the parameters λ and μ for both the sigmoid strictly increasing function $v_s(x; \lambda, \mu) = 1/(1 + e^{-\lambda(x-\mu)})$ and the strictly decreasing function $\theta(x; \mu) = 2\mu - x$ according to the following rationale. Given the dynamics of our dispensing system in practice it is required an "early warning" signal at $v = 1486$ $[lt]$. It is already known that an IN induced from a uniform distribution has an isosceles triangular shape [35]. Therefore, in line with our assumptions above, a population of measurements with an average equal to L was repesented by an isosceles triangular IN $V = V(h) = [a_h, b_h]$, $h \in [0, 1]$ with support $[L-3, L+3]$. There follows $a_h = 3h + (L-3)$ as well as $b_h = (L+3) - 3h$ for $h \in [0, 1]$. In conclusion, inclusion measure $\sigma_{\curlyvee}(V, V_{ref})$ equals

$$\sigma^V_{\curlyvee}(V, V_{ref}) =$$
$$= \int_0^1 \frac{v_s(\theta(1496.4)) + v_s(1503.6)}{v_s(\theta((3h+1483) \wedge 1496.4)) + v_s((1489-3h) \vee 1503.6)} dh =$$
$$= \sigma_{\curlyvee}(\lambda, \mu).$$

Fig.8 displays the degree of membership $\sigma^V_{\curlyvee}(V, V_{ref})$ regarding an isosceles triangular IN V with support $[1483, 1489]$ located at $L = 1486$, as a function of the sigmoid function $v_s(v; \lambda, \mu)$ parameters λ and μ. It is preferable to select a pair (λ, μ) of parameter values that results in a *small* value of $\sigma^V_{\curlyvee}(V, V_{ref})$ so as to secure an easily detectable "early warning" signal. Fig.8 indicates that smaller values of function $\sigma^V_{\curlyvee}(\lambda, \mu)$ are attained for both $\mu \simeq 1492$ and large values of λ. In fact, it can be (easily) shown analytically that $\lim\limits_{\lambda \to +\infty} \sigma^V_{\curlyvee}(\lambda, \mu = 1492) = 0.5$. We decided to use $\lambda = 1$ so as to retain the typical sigmoid function shape. Furthermore, using a steepest descent method we computed $\mu = 1492.270$ resulting in an acceptable optimal (minimum) value $\sigma^V_{\curlyvee}(\lambda = 1, \mu = 1492.270) = 0.516$. We point out that, in previous works, optimal parameter estimation "of scale" has been pursued using stochastic search techniques such as genetic algorithms [36], [42], [58].

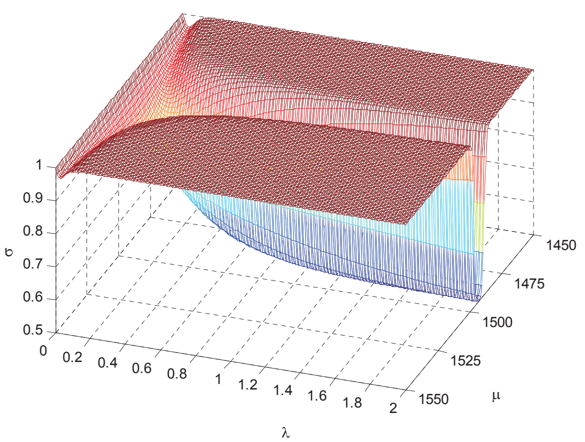


Fig. 8. The 2-dimensional curve above shows the degree of membership $\sigma^V_{\curlyvee}(V, V_{ref})$, regarding an isosceles triangular IN V with support $[1486 - 3, 1486 + 3]$ and $V_{ref} = V_{ref}(h) = [1496.4, 1503.6]$ for $h \in [0, 1]$, as a function of the sigmoid function $v_s(v; \lambda, \mu)$ parameters λ and μ.

Fig.9 displays INs V_1, V_2 and V_3 induced by algorithm CALCIN from the measurements of a liquid being dispensed into the mixing tank, whereas IN $V_{ref}(h) = [V_0 - \Delta V, V_0 + \Delta V], h \in [0, 1]$ is the *reference volume* with $V_0 = 1500$ $[lt]$ and $\Delta V = 3.6$ $[lt]$. IN V_1 was induced first, followed by IN V_2, the latter in turn was followed by IN V_3. In general, not only the peak of an IN but also its shape changes with time because a different distribution of samples is obtained at a different time. Fig. 9 also displays the strictly increasing (sigmoid) function $v_s(v; \lambda, \mu) = \frac{1}{1+e^{-\lambda(v-\mu)}}$ employed here with the optimally estimated parameter values $\lambda = 1$ and $\mu = 1492.270$. In all cases, the strictly decreasing function $\theta(v; \mu) = 2\mu - v = 2984.54 - v$ was employed.

Fig.10 (a) and (b) display inclusion measure $\sigma^V_{\curlywedge}(V, V_{ref})$ and $\sigma^V_{\curlyvee}(V, V_{ref})$, respectively, furthermore Fig.10(c) displays the result of applying Zadeh's compositional rule of inference (CRI) $v = V \circ (V_{ref} \Rightarrow$ "*stop dispensing*") versus the dispensed liquid volume v over the range $[1480, 1520]$. Fig.10 demonstrates that either function $\sigma^V_{\curlywedge}(V, V_{ref})$ or the CRI

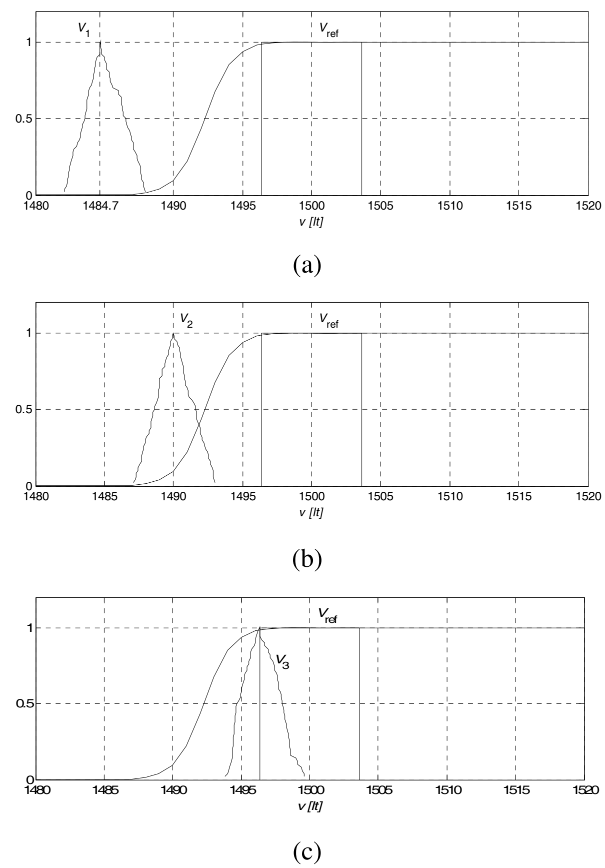


Fig. 9. (a) IN V_1 was induced first, followed by (b) IN V_2, followed by (c) IN V_3. The reference volume IN V_{ref} is displayed in all figures as well as the optimally estimated, as explained in the text, sigmoid function $v_s(v; 1, 1492.270) = 1/(1 + e^{-(v-1492.270)})$. In (c) the average μ_V of a population of measurements regarding a liquid volume equals $\mu_V = 1496.4$.

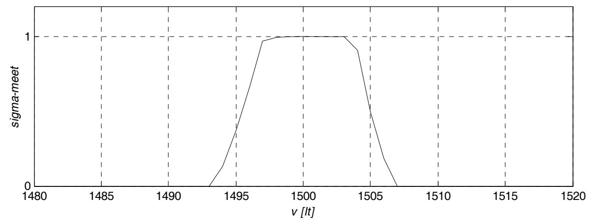
signify more "crisply" than function $\sigma^V_{\curlyvee}(V, V_{ref})$ the order relation "$V \preceq V_{ref}$" in the sense that either $\sigma^V_{\curlywedge}(V, V_{ref})$ or the result by CRI rises from 0 all the way to 1, whereas $\sigma^V_{\curlyvee}(V, V_{ref})$ rises only from (slightly over) 0.5 to 1. Nevertheless, only the inclusion measure function $\sigma^V_{\curlyvee}(V, V_{ref})$ can warn as early as at around $v = 1486$ that we approach the reference volume V_{ref}. Hence, the inclusion measure $\sigma^V_{\curlyvee}(V, V_{ref})$ appears to be a better decision-making instrument in practice than either inclusion measure $\sigma^V_{\curlywedge}(V, V_{ref})$ or the CRI towards *stop dispensing* within specifications.

C. Comparative Experimental Results

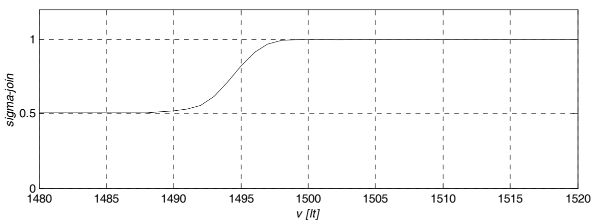
A standard practice in the industry for dealing with a population of measurements is to replace it with its first-order data statistic, namely the population average. Therefore, we considered the following alternative decision-making rule for dispensing a liquid to the mixing tank.

Rule R_1 : IF $|\mu_V - V_0| \leq \Delta V$ THEN *stop dispensing*,

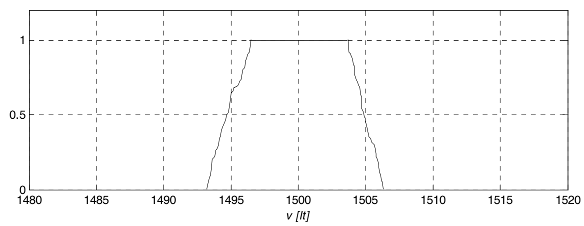
where μ_V is the average of a population of measurements regarding the volume of a liquid being dispensed, whereas both V_0 and ΔV have been defined above.



(a)



(b)



(c)

Fig. 10. Using the optimal parameter estimated functions $v_s(v; \lambda = 1, \mu = 1492.270)$ and $\theta(v; \mu = 1492.270)$ we computed (a) Inclusion measure $\sigma_{\curlywedge}^{V}(V, V_{ref})$, and (b) Inclusion measure $\sigma_{\dot{\curlyvee}}^{V}(V, V_{ref})$, all versus the dispensed liquid volume $v \in [1480, 1520]$. Alternatively, (c) presents the corresponding result by Zadeh's compositional rule of inference (CRI) versus the dispensed liquid volume $v \in [1480, 1520]$.

Fig.9(c) illustrates how rule R_1 can be activated while, at the same time, the previous rule R_0 remains inactive for $T = 1$. The practical problem in this case is that liquid dispensing stops while the actual volume of the dispensed liquid might be less than $|V_0 - \Delta V| = |1500 - 3.6| = 1496.4$; hence, the final industrial product might be outside specifications. The aforementioned problem is dubbed here "false triggering" and it can be resolved in Fig.9(c) using rule R_0 with either $\sigma_{\curlywedge}(.,.)$ or $\sigma_{\dot{\curlyvee}}(.,.)$ and $T = 1$.

It might be thought that, under the (numerical) assumptions of Fig.9(c), "false triggering" can be avoided using the following alternative rule

Rule R_2 : IF $\mu_V = V_0$ THEN *stop dispensing*.

However, rule R_2 might not be able to deal with another problem; the latter occurs when the support "$2 \cdot \Delta V$" of IN V_{ref} is smaller than the support of IN V, furthermore the problem exacerbates when IN V is skewed thus deteriorating performance as detailed in [35].

In a series of computational experiments, using rule R_0 with either inclusion measure $\sigma_{\curlywedge}^{V}(.,.)$ or $\sigma_{\dot{\curlyvee}}^{V}(.,.)$ and a user-defined threshold $T = 0.93$, it turns out that rule R_0 clearly maximizes the probability of *stop dispensing* a liquid within specifications. The latter probability corresponded to the portion of IN V over the interval $[V_0 - \Delta V, V_0 + \Delta V]$ at the very moment liquid dispensing stops due to the activation of the rule in use. Our explanation for the superior performance of an inclusion measure (in rule R_0) is that $\sigma_{\curlywedge}^{V}(.,.)$ as well as $\sigma_{\dot{\curlyvee}}^{V}(.,.)$ engage all-order data statistics, whereas an alternative rule typically engages fewer (user-defined) data statistics such as the corresponding average and standard deviation, i.e. first- and second- order data statistics, respectively.

This preliminary industrial dispensing application was meant to demonstrate the practical applicability of our proposed techniques rather than to analyze their efficiency. A comparative study regarding the efficiency of our proposed techniques including potential improvements is a topic for future research.

VI. Discussion & Conclusion

The thrust of this paper is the introduction of novel perspectives as well as sound mathematical results, including theorems 3.13 and 3.17, towards a "principled" (in the sense of satisfying the properties **C1-C3** of Definition 3.1) extension of a Fuzzy Inference System (FIS) involving arbitrary (fuzzy number) membership function shapes. In particular, this work has introduced a number of FIS extensions regarding the *matching degree* (or satisfaction degree, or firing degree) part of a FIS. This has been achieved by studying the lattice $(\mathbb{I}, \subseteq)$ of conventional intervals on the line of real numbers followed by a constructive study of the lattice $(\mathbb{F}, \preceq)$ of Intervals' Numbers (INs). Lattice $(\mathbb{F}, \preceq)$ was shown to be isomorphic to the lattice $(\mathbb{F}', \leq)$ of fuzzy intervals. Two inclusion measures $\sigma_{\cap}^{V}(.,.)$ and $\sigma_{\dot{\cup}}^{V}(.,.)$ were introduced on $(\mathbb{I}, \subseteq)$ giving rise to inclusion measures $\sigma_{\curlywedge}^{V}(.,.)$ and $\sigma_{\dot{\curlyvee}}^{V}(.,.)$, respectively, on $(\mathbb{F}, \preceq)$.

Based on theorems 3.13 and 3.17 we showed that inclusion measures are widely (though *implicitly*) used by traditional FISs. Examples 4.1-4.4 indicated that an explicit employment of an inclusion measure (σ) may result in substantial benefits including: 1) Accommodation of granular FIS inputs, 2) Employment of sparse FIS fuzzy rule bases, and 3) Introduction of tunable nonlinearities globally, rather than locally, via parametric *length functions*, while retaining traditional FIS semantics.

Appendix

Theorem 2.3

Proof:

$F \preceq G \Leftrightarrow (\forall h \in [0,1] : F(h) \subseteq G(h))$ by definition.

Suppose $(\forall h \in [0,1] : F(h) \subseteq G(h))$ holds. Take any $x \in \mathbb{L}$ and let $h = m_F(x)$. Then $x \in F(h) \subseteq G(h) \Rightarrow x \in G(h) \Rightarrow m_G(x) \geq h = m_F(x)$.

Suppose that $(\forall x \in \mathbb{L} : m_F(x) \leq m_G(x))$ holds. Take any $h \in [0,1]$. If $F(h)$ is empty, then $F(h) \subseteq G(h)$. If $F(h)$ is not empty, take any $x \in F(h)$. Then $h \leq m_F(x) \leq m_G(x)$ and so $x \in G(h)$. Hence, $F(h) \subseteq G(h)$. ■

Theorem 3.2

Proof:

(i) We first prove: $\sigma(x,y) = 1 \Rightarrow x \sqsubseteq y$. This is simply the contrapositive of **C2**.
(ii) Next we prove: $x \sqsubseteq y \Rightarrow \sigma(x,y) = 1$. To do this, replace in **C3** y with x and z with y, to get $x \sqsubseteq y \Rightarrow \sigma(x,x) \leq \sigma(x,y)$; but $\sigma(x,x) = 1$ (from **C1**) and $\sigma(x,y) \leq 1$ (since $\sigma : \mathbb{X} \times \mathbb{X} \to [0,1]$), hence $x \sqsubseteq y \Rightarrow \sigma(x,y) = 1$.
(iii) Now we prove $x \sqcap y \sqsubset x \Rightarrow \sigma(x,y) < 1$. As already proved, $\sigma(x,y) = 1 \Rightarrow x \sqsubseteq y \Rightarrow x \sqcap y = x$. Using the contrapositive of this, we have $x \sqcap y \neq x \Rightarrow x \not\sqsubseteq y \Rightarrow \sigma(x,y) < 1$.
(iv) Finally we prove $\sigma(x,y) < 1 \Rightarrow x \sqcap y \sqsubset x$. Choose x and y such that $\sigma(x,y) < 1$ and assume $x \sqcap y \not\sqsubset x$; then clearly $x \sqcap y = x$ and so $x \sqsubseteq y$. But then, from (ii) we get $\sigma(x,y) = 1$ which contradicts $\sigma(x,y) < 1$. Hence $\sigma(x,y) < 1 \Rightarrow x \sqcap y \sqsubset x$. ■

Lemma 3.4

Proof:

Suppose first that $\emptyset = O = A \subset B = [b_1, b_2]$; then $V(A) = 0 < v(\theta(b_1)) + v(b_2) = V(B)$. If, on the other hand, $O \subset A = [a_1, a_2] \subset B = [b_1, b_2]$, then

$$\text{either } b_1 \leq a_1 \leq a_2 < b_2 \quad \text{or} \quad b_1 < a_1 \leq a_2 \leq b_2$$
$$\text{or} \quad b_1 < a_1 \leq a_2 < b_2.$$

We will only consider the first case (the others are treated similarly). If $b_1 \leq a_1 \leq a_2 < b_2$, we have $\theta(a_1) \leq \theta(b_1)$ and so $v(\theta(a_1)) \leq v(\theta(b_1))$; also $v(a_2) < v(b_2)$. And so

$$V(A) = v(\theta(a_1)) + v(a_2) < v(\theta(b_1)) + v(b_2) = V(B).$$

■

Lemma 3.5

Proof:

To prove (3), let us distinguish two cases.
i. If $A \subseteq C$ then we have

$$C = A \dot\cup C \Rightarrow V\left(A \dot\cup C\right) - V(C) = 0$$
$$B \subseteq A \dot\cup B \Rightarrow V\left(A \dot\cup B\right) - V(B) \geq 0$$

which proves (3).
ii. If $A \not\subseteq C$ then also $A = [a_1, a_2] \supset O$. Either $a_1 < c_1 \leq b_1$ or $b_2 \leq c_2 < a_2$ (or both). We examine the two subcases separately.
ii.1. If $a_1 < c_1 \leq b_1$ we have

$$A \dot\cup C = [a_1, a_2 \vee c_2], \qquad C = [c_1, c_2]$$
$$V\left(A \dot\cup C\right) - V(C) =$$
$$= v(\theta(a_1)) + v(a_2 \vee c_2) - v(\theta(c_1)) - v(c_2)$$

and

$$A \dot\cup B = [a_1, a_2 \vee b_2], \qquad B = [b_1, b_2]$$
$$V\left(A \dot\cup B\right) - V(B) =$$
$$= v(\theta(a_1)) + v(a_2 \vee b_2) - v(\theta(b_1)) - v(b_2).$$

So, to test the validity of (3), we must compare

$$v(a_2 \vee c_2) - v(\theta(c_1)) - v(c_2) \qquad \text{and}$$
$$v(a_2 \vee b_2) - v(\theta(b_1)) - v(b_2).$$

Now

$$c_1 \leq b_1 \Rightarrow \theta(b_1) \leq \theta(c_1) \Rightarrow v(\theta(b_1)) \leq v(\theta(c_1)) \Rightarrow$$
$$\Rightarrow -v(\theta(c_1)) \leq -v(\theta(b_1)).$$

Also, for the relative position of a_2, b_2, c_2 we have three possibilities
ii.1.1 If $a_2 \leq b_2 \leq c_2$ then

$$v(a_2 \vee c_2) - v(c_2) = v(c_2) - v(c_2) = v(b_2) - v(b_2) =$$
$$= v(a_2 \vee b_2) - v(b_2).$$

ii.1.2 If $b_2 \leq a_2 \leq c_2$ then

$$v(a_2 \vee c_2) - v(c_2) = v(c_2) - v(c_2) \leq v(a_2) - v(b_2) =$$
$$= v(a_2 \vee b_2) - v(b_2).$$

ii.1.3 If $b_2 \leq c_2 \leq a_2$ then

$$v(a_2 \vee c_2) - v(c_2) = v(a_2) - v(c_2) \leq v(a_2) - v(b_2) =$$
$$= v(a_2 \vee b_2) - v(b_2).$$

Hence (3) holds in this case.
ii.2. The treatment of the case $b_2 \leq c_2 < a_2$ is similar to that of **ii.1** and hence is omitted. It turns out that (3) holds in this case too.

Hence (3) holds in every case and the proof of the Lemma is complete. ■

Theorem 3.7

Proof:

First let us verify that Properties **C1-C3** hold for $\sigma_{\cap}^{V}$.
i. If $A = O$ then $\sigma_{\cap}^{V}(A, A) = 1$. If $A \supset O$ then $\sigma_{\cap}^{V}(A, A) = \frac{V(A \cap A)}{V(A)} = 1$ and **C1** holds.
ii. Assume $A \not\subseteq B$. Then $O \subset A$, $A \cap B \subset A$ and $V(A \cap B) < V(A)$. Hence $\sigma_{\cap}^{V}(A, B) = \frac{V(A \cap B)}{V(A)} < 1$ and **C2** holds.
iii. Assume $B \subseteq C$; then we also have $A \cap B \subseteq A \cap C$ and $V(A \cap B) \leq V(A \cap C)$. Now consider two cases.
iii.1 First, suppose $A = O$. Then $\sigma_{\cap}^{V}(A, B) = 1 = \sigma_{\cap}^{V}(A, C)$.
iii.2 Second, suppose $A \supset O$. Then $\sigma_{\cap}^{V}(A, B) = \frac{V(A \cap B)}{V(A)} \leq \frac{V(A \cap C)}{V(A)} = \sigma_{\cap}^{V}(A, C)$.
Hence **C3** holds.

Next let us verify Properties **C1-C3** for $\sigma_{\dot\cup}^{V}$.
i. If $A = O$ then $\sigma_{\dot\cup}^{V}(A, A) = 1$. If $A \supset O$ then $\sigma_{\dot\cup}^{V}(A, A) = \frac{V(A)}{V\left(A \dot\cup A\right)} = 1$ and **C1** holds.
ii. Assume $A \not\subseteq B$. Then $O \subset A \subset A \dot\cup B$ and $A \cap B \subset A$. Also, either $B \subset A$ or *both* $A \cap B \subset A \subset A \dot\cup B$ and $A \cap B \subset B \subset A \dot\cup B$.
ii.1 If $B \subset A$ then also $A = A \dot\cup B$, $0 < V\left(A \dot\cup B\right)$ and $V(B) < V(A) = V\left(A \dot\cup B\right)$, hence $\sigma_{\dot\cup}^{V}(A, B) = \frac{V(B)}{V\left(A \dot\cup B\right)} < 1$.
ii.2 If $A \cap B \subset A \subset A \dot\cup B$ and $A \cap B \subset B \subset A \dot\cup B$ then also $V(B) < V\left(A \dot\cup B\right)$, hence $\sigma_{\dot\cup}^{V}(A, B) = \frac{V(B)}{V\left(A \dot\cup B\right)} < 1$.
In both cases **C2** holds.
iii. Assume $B \subseteq C$.

iii.1. If $B = O$, we distinguish two subcases.
iii.1.1. If also $A = O$, $\sigma^V_{\dot\cup}(A, B) = 1$. But also $C = A \dot\cup C$ and hence (for both $C = O$ and $C \supset O$): $\sigma^V_{\dot\cup}(A, C) = 1 = \sigma^V_{\dot\cup}(A, B)$.
iii.1.2. If $A \supset O$, then $V(A) > 0$. Also $O \subset A \dot\cup B \subseteq A \dot\cup C \Rightarrow 0 < V\left(A \dot\cup B\right) \leq V\left(A \dot\cup C\right)$. Hence

$$\sigma^V_{\dot\cup}(A, B) = \frac{V(B)}{V\left(A \dot\cup B\right)} = 0 \leq \frac{V(C)}{V\left(A \dot\cup C\right)} = \sigma^V_{\dot\cup}(A, C)$$

iii.2. If $B \supset O$ then $V\left(A \dot\cup B\right) > 0$ and, using Lemma 3.5, we have

$$\begin{aligned}
&V\left(A \dot\cup C\right) - V(C) \leq V\left(A \dot\cup B\right) - V(B) \Rightarrow \\
&V\left(A \dot\cup C\right) \leq V(C) - V(B) + V\left(A \dot\cup B\right) \Rightarrow \\
&V(B) \cdot V\left(A \dot\cup C\right) \leq V(B) \cdot \left(V(C) - V(B) + V\left(A \dot\cup B\right)\right) \Rightarrow \\
&V(B) \cdot V\left(A \dot\cup C\right) \leq V(B) \cdot (V(C) - V(B)) + \\
&+V(B) \cdot V\left(A \dot\cup B\right) \Rightarrow \\
&V(B) \cdot V\left(A \dot\cup C\right) \leq V\left(A \dot\cup B\right) \cdot (V(C) - V(B)) + \\
&+V(B) \cdot V\left(A \dot\cup B\right) \Rightarrow \\
&V(B) \cdot V\left(A \dot\cup C\right) \leq V\left(A \dot\cup B\right) \cdot V(C) \Rightarrow \\
&\frac{V(B)}{V\left(A \dot\cup B\right)} \leq \frac{V(C)}{V\left(A \dot\cup C\right)} \Rightarrow \\
&\sigma^V_{\dot\cup}(A, B) \leq \sigma^V_{\dot\cup}(A, C).
\end{aligned}$$

In both cases **C3** holds. ■

Theorem 3.12

Proof:

We will only verify Properties **C1-C3** for $\sigma^V_{\curlywedge}$ (the case of $\sigma^V_{\curlyvee}$ can be proved similarly).

C1 We want to prove: for any $F \in \mathbb{F}$, we have $\sigma^V_{\curlywedge}(F, F) = 1$. We partition $[0, 1]$ into two sets A and B, where $A = \{h : F(h) = O\}$, $B = \{h : F(h) \supset O\}$. Then, as already seen, $\sigma^V_{\cap}(F(h), F(h)) = 1$ for all $h \in A$; and also $\sigma^V_{\cap}(F(h), F(h)) = \frac{V(F(h) \cap F(h))}{V(F(h))} = 1$ for all $h \in B$. And so

$$\sigma^V_{\curlywedge}(F, F) = \int_0^1 \sigma^V_{\cap}(F(h), F(h))dh =$$

$$= \int_A \sigma^V_{\cap}(F(h), F(h))dh + \int_B \sigma^V_{\cap}(F(h), F(h))dh =$$

$$= \int_A 1dh + \int_B 1dh = \int_0^1 1dh = 1.$$

C2 We want to prove: for any $F, G \in \mathbb{F}$ such that $F \not\preceq G$ we have $\sigma^V_{\curlywedge}(F, G) < 1$. Note that $F \not\preceq G \Rightarrow F \curlywedge G \prec F$. Then, according to Corollary 2.4 we have:
(a) $\forall x : m_{F \curlywedge G}(x) = m_F(x) \wedge m_G(x) \leq m_F(x)$ and
(b) $\exists x_0 : m_{F \curlywedge G}(x_0) = m_F(x_0) \wedge m_G(x_0) = h_1 < h_2 = m_F(x_0)$. Then, we have

$$\begin{aligned}
\forall h \in (h_1, h_2] : m_F(x_0) \wedge m_G(x_0) = h_1 < h \leq h_2 = \\
= m_F(x_0) \Rightarrow \\
\forall h \in (h_1, h_2] : x_0 \notin (F \curlywedge G)(h) \text{ and } x_0 \in F(h) \Rightarrow \\
\forall h \in (h_1, h_2] : (F \curlywedge G)(h) \subset F(h) \Rightarrow \\
\forall h \in (h_1, h_2] : \sigma^V_{\cap}(F(h), G(h)) < 1.
\end{aligned}$$

Hence,

$$\sigma^V_{\curlywedge}(F, G) = \int_0^1 \sigma^V_{\cap}(F(h), G(h))dh =$$

$$= \int_{(h_1, h_2]} \sigma^V_{\cap}(F(h), G(h))dh +$$

$$+ \int_{[0,1] \setminus (h_1, h_2]} \sigma^V_{\cap}(F(h), G(h))dh <$$

$$< (h_2 - h_1) + (1 - (h_2 - h_1)) = 1.$$

C3 We want to prove: for any $F, G, A \in \mathbb{F}$ with $F \preceq G$, we have $\sigma^V_{\curlywedge}(A, F) \leq \sigma^V_{\curlywedge}(A, G)$. Indeed, for all $h \in [0, 1]$ we have $F(h) \subseteq G(h)$ and so $\sigma^V_{\cap}(A(h), F(h)) \leq \sigma^V_{\cap}(A(h), G(h))$ which means

$$\sigma^V_{\curlywedge}(A, F) = \int_0^1 \sigma^V_{\cap}(A(h), F(h))dh \leq$$

$$\leq \int_0^1 \sigma^V_{\cap}(A(h), G(h))dh = \sigma^V_{\curlywedge}(A, G).$$

■

Theorem 3.13

Proof:

Take any $h \in [0, 1]$. We have $x_0 \in F(h) \Leftrightarrow m_{\widetilde{F}}(x_0) \geq h$ or, equivalently, $x_0 \notin F(h) \Leftrightarrow m_{\widetilde{F}}(x_0) < h$. Now:

$$\begin{aligned}
x_0 \in F(h) \Rightarrow \sigma^V_{\cap}([x_0, x_0], F(h)) = \frac{V([x_0, x_0] \cap F(h))}{V([x_0, x_0])} = \\
= \frac{V([x_0, x_0])}{V([x_0, x_0])} = 1, \\
x_0 \notin F(h) \Rightarrow \sigma^V_{\cap}([x_0, x_0], F(h)) = \frac{V([x_0, x_0] \cap F(h))}{V([x_0, x_0])} = \\
= \frac{V(O)}{V([x_0, x_0])} = 0.
\end{aligned}$$

Define the sets

$$\begin{aligned}
A = \{h : x_0 \in F(h)\} = \{h : h \leq m_{\widetilde{F}}(x_0)\} = [0, m_{\widetilde{E}}(x_0)], \\
B = \{h : x_0 \notin F(h)\} = \{h : h > m_{\widetilde{F}}(x_0)\} = (m_{\widetilde{E}}(x_0), 1].
\end{aligned}$$

Then

$$\sigma^V_{\curlywedge}(X_0, F) = \int_0^1 \sigma^V_{\cap}(X_0, F(h))dh =$$

$$= \int_0^{m_{\widetilde{F}}(x_0)} 1dh + \int_{m_{\widetilde{F}}(x_0)}^{1} 0dh = m_{\widetilde{F}}(x_0).$$

and the proof is complete. ■

Theorem 3.17

Proof:

We just check that **C1** – **C3** of Definition 3.1 are satisfied.

C1 For any $\mathbf{x} \in \mathbb{X}$, $\sigma_\wedge(\mathbf{x},\mathbf{x}) = \min_{i\in\{1,\dots,N\}} \sigma_i(x_i,x_i) = 1$ and $\sigma_\Pi(\mathbf{x},\mathbf{x}) = \prod_{i=1}^{N}\sigma_i(x_i,x_i) = 1$.

C2 Take any $\mathbf{x},\mathbf{y} \in \mathbb{X}$ such that $\mathbf{x} \not\sqsubseteq \mathbf{y}$. Then also

$$\mathbf{x} \sqcap \mathbf{y} \sqsubset \mathbf{x} \Rightarrow (\exists n \in \{1,\dots,N\} : x_n \sqcap_n y_n \sqsubset_n x_n) \Rightarrow$$
$$\Rightarrow \sigma_n(x_n,y_n) < 1.$$

Hence, $\sigma_\wedge(\mathbf{x},\mathbf{y}) = \min_{i\in\{1,\dots,N\}} \sigma_i(x_i,y_i) < 1$ and $\sigma_\Pi(\mathbf{x},\mathbf{y}) = \prod_{i=1}^{N}\sigma_i(x_i,y_i) < 1$.

C3 Take any $\mathbf{u},\mathbf{w} \in \mathbb{X}$ such that $\mathbf{u} \sqsubseteq \mathbf{w}$. More specifically, let $(u_1,\dots,u_N) = \mathbf{u} \sqsubseteq \mathbf{w} = (w_1,\dots,w_N)$. Now take any $\mathbf{x} = (x_1,\dots,x_N) \in \mathbb{X}$. We have

$$(\forall i \in \{1,\dots,N\} : u_i \sqsubseteq_i w_i) \Rightarrow (\forall i \in \{1,\dots,N\} :$$
$$\sigma_i(x_i,u_i) \le \sigma_i(x_i,w_i)).$$

Hence,

$$\min_{i\in\{1,\dots,N\}} \sigma_i(x_i,u_i) \le \min_{i\in\{1,\dots,N\}} \sigma_i(x_i,w_i) \Rightarrow$$
$$\Rightarrow \sigma_\wedge(\mathbf{x},\mathbf{u}) \le \sigma_\wedge(\mathbf{x},\mathbf{w})$$

and

$$\prod_{i=1}^{N}\sigma_i(x_i,u_i) \le \prod_{i=1}^{N}\sigma_i(x_i,w_i) \Rightarrow$$
$$\Rightarrow \sigma_\Pi(\mathbf{x},\mathbf{u}) \le \sigma_\Pi(\mathbf{x},\mathbf{w}).$$

The proof is complete.

■

ACKNOWLEDGMENT

This work has been supported, in part, by voucher no. 16106672-01-000115 of the Greek government's "Vouchers for Small & Medium Enterprises" action as well as by the European Union (Social Fund) and Greek national resources under the framework of the "Archimedes III: Funding of Research Groups in TEI of Athens" project of the "Education & Lifelong Learning" Operational Programme.

REFERENCES

[1] W. Bandler and L. Kohout, "Fuzzy power sets and fuzzy implication operators," *Fuzzy Sets and Systems*, vol. 4, no. 1, pp. 13–30, 1980.

[2] R. Belohlavek, *Fuzzy Relational Systems: Foundations and Principles*. Heidelberg, Germany: Springer, 2002.

[3] R. Belohlavek and G.J. Klir, eds., *Concepts and Fuzzy Logic*. Cumberland, RI: The MIT Press, 2011.

[4] G. Birkhoff, *Lattice Theory*. Providence, RI: American Mathematical Society, 1967, Colloquium Publications, vol. 25.

[5] R. Boukezzoula, S. Galichet, and A. Bisserier, "A midpoint-radius approach to regression with interval data," *International Journal of Approximate Reasoning*, vol. 52, no. 9, pp. 1257–1271, 2011.

[6] P. Burillo, N. Frago, and R. Fuentes, "Inclusion grade and fuzzy implication operators," *Fuzzy Sets and Systems*, vol. 114, no. 3, 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 Approx. Reasoning*, vol. 23, no. 3, pp. 137–209, 2000.

[8] S.-M. Chen, W.-C. Hsin, S.-W. Yang, and Y.-C. Chang, "Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on the slopes fuzzy sets," *Expert Systems with Applications*, vol. 39, no. 15, pp. 11961–11969, 2012.

[9] O. Cordón, "A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems," *International Journal of Approximate Reasoning*, vol. 52, no. 6, pp. 894–913, 2011.

[10] C. Cornelis and E. Kerre, "Inclusion measures in intuitionistic fuzzy set theory," in *Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (Lecture Notes in Artificial Intelligence Series 2711), T.D. Nielsen and N.L. Zhang, Eds. Berlin, Germany: Springer, 2003, pp. 345–356.

[11] C. Cornelis, C. Van der Donck, and E. Kerre, "Sinha–Dougherty approach to the fuzzification of set inclusion revisited," *Fuzzy Sets and Systems*, 134, no. 2, pp. 283-295, 2003.

[12] G. Deschrijver and E.E. Kerre, "On the relationship between some extensions of fuzzy set theory," *Fuzzy Sets and Systems*, vol. 133, no. 2, pp. 227–235, 2003.

[13] G. Deschrijver and E.E. Kerre, "On the position of intuitionistic fuzzy set theory in the framework of theories modelling imprecision," *Information Sciences*, vol. 177, no. 8, pp. 1860–1866, 2007.

[14] E. Esmi, P. Sussner, M.E. Valle, F. Sakuray, and L. Barros, "Fuzzy associative memories based on subsethood and similarity measures with applications to speaker identification," in *Hybrid Artificial Intelligent Systems* (Lecture Notes in Computer Science Series 7209), E. Corchado et al., Eds. Berlin, Germany: Springer-Verlag, 2012, vol. II, pp. 479–490.

[15] J. Fan, W. Xie, and J. Pei, "Subsethood measure: new definitions," *Fuzzy Sets and Systems*, vol. 106, no. 2, pp. 201–209, 1999.

[16] L. Foulloy and S. Galichet, "Fuzzy control with fuzzy inputs," *IEEE Transactions on Fuzzy Systems*, vol. 11, no. 4, pp. 437–449, 2003.

[17] B. Ganter and R. Wille, *Formal Concept Analysis*. Heidelberg, Germany: Springer, 1999.

[18] J.A. Goguen, "L-fuzzy sets," *Journal of Mathematical Analysis and Applications*, vol. 18, no. 1, pp. 145–174, 1967.

[19] M. Grabisch, "Belief functions on lattices," *International Journal of Intelligent Systems*, vol. 24, no. 1, pp. 76–95, 2009.

[20] M. Graña, "Lattice computing in hybrid intelligent systems," in *Proceedings of the 12th Intl. Conf. on Hybrid Intelligent Systems (HIS)*, 4-7 Dec. 2012, pp. 1–5.

[21] M. Graña (Ed.), Special issue on: Lattice computing and natural computing, *Neurocomputing*, vol. 72, no. 10-12, pp. 2065–2066, 2009.

[22] M. Graña and A.I. Gonzalez-Acuña, "Learning parsimonious dendritic classifiers," *Neurocomputing*, vol. 109, pp. 3–8, 2013.

[23] P. Grzegorzewski, "On possible and necessary inclusion of intuitionistic fuzzy sets," *Info. Sciences*, vol.181, no. 2, pp. 342–350, 2011.

[24] S. Guillaume, "Designing fuzzy inference systems from data: an interpretability-oriented review," *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 3, pp. 426–443, 2001.

[25] H.J.A.M. Heijmans and R. Keshet, "Inf-semilattice approach to self-dual morphology," *Journal of Mathematical Imaging and Vision*, vol. 17, no. 1, pp. 55–80, 2002.

[26] Z. Huang and Q. Shen, "Fuzzy interpolation and extrapolation: a practical approach," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 1, pp. 13–28, 2008.

[27] C.-M. Hwang, M.-S. Yang, and W.-L. Hung, "On similarity, inclusion measure and entropy between type-2 fuzzy sets," *Intl. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 20, no. 3, pp. 433–449, 2012.

[28] J. Jiménez, S. Montes, B. Šešelja, and A. Tepavčević, "Lattice-valued approach to closed sets under fuzzy relations: theory and applications," *Computers and Mathematics with Applications*, vol. 62, no. 10, pp. 3729–3740, 2011.

[29] V.G. Kaburlasos, *Adaptive Resonance Theory With Supervised Learning and Large Database Applications*. Ph.D. Dissertation, Univ. Nevada, Reno, 1992, USA Library of Congress-Copyright Office.

[30] V.G. Kaburlasos, "FINs: lattice theoretic tools for improving prediction of sugar production from populations of measurements," *IEEE Transactions on Systems, Man & Cybernetics - Part B*, vol. 34, no. 2, pp. 1017–1030, 2004.

[31] V.G. Kaburlasos, *Towards a Unified Modeling and Knowledge-Representation Based on Lattice Theory*. Heidelberg, Germany: Springer, 2006, ser. Studies in Computational Intelligence, vol. 27.
[32] V.G. Kaburlasos (Ed.), Special issue on: Information engineering applications based on lattices, *Information Sciences*, vol. 181, no. 10, pp. 1771–1773, 2011.
[33] V.G. Kaburlasos and A. Kehagias, "Novel fuzzy inference system (FIS) analysis and design based on lattice theory. part I: working principles," *International Journal of General Systems*, vol. 35, no. 1, pp. 45–67, 2006.
[34] V.G. Kaburlasos and A. Kehagias, "Novel fuzzy inference system (FIS) analysis and design based on lattice theory," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 2, pp. 243–260, 2007.
[35] V.G. Kaburlasos and T. Pachidis, "A lattice-computing ensemble for reasoning based on formal fusion of disparate data types, and an industrial dispensing application," *Information Fusion*, in press
[36] V.G. Kaburlasos and S.E. Papadakis, "Granular self-organizing map (grSOM) for structure identification," *Neural Networks*, vol. 19, no. 5, pp. 623–643, 2006.
[37] V.G. Kaburlasos and V. Petridis, "Fuzzy lattice neurocomputing (FLN): a novel connectionist scheme for versatile learning and decision making by clustering," *International Journal of Computers and Their Applications*, vol. 4, no. 3, pp. 31–43, 1997.
[38] V.G. Kaburlasos and V. Petridis, "Fuzzy lattice neurocomputing (FLN) models," *Neural Networks*, vol. 13, no. 10, pp. 1145–1169, 2000.
[39] V.G. Kaburlasos and V. Petridis, "Learning and decision-making in the framework of fuzzy lattices," in *New Learning Paradigms in Soft Computing* (Studies in Fuzziness and Soft Computing Series 84), L.C. Jain and J. Kacprzyk, Eds. Heidelberg, Germany: Physica-Verlag, 2002, pp. 55–96.
[40] V.G. Kaburlasos and G.X. Ritter (Eds.), *Computational Intelligence Based on Lattice Theory*. Heidelberg, Germany: Springer, 2007, ser. Studies in Computational Intelligence, vol. 67.
[41] V.G. Kaburlasos, I.N. Athanasiadis, and P.A. Mitkas, "Fuzzy lattice reasoning (FLR) classifier and its application for ambient ozone estimation," *International Journal of Approximate Reasoning*, vol. 45, no. 1, pp. 152–188, 2007.
[42] V.G. Kaburlasos, S.E. Papadakis, and A. Amanatiadis, "Binary image 2D shape learning and recognition based on lattice-computing (LC) techniques," *Journal of Mathematical Imaging and Vision*, vol. 42, no. 2-3, pp. 118–133, 2012.
[43] V.G. Kaburlasos, S.E. Papadakis, and G.A. Papakostas, "Lattice computing extension of the FAM neural classifier for human facial expression recognition," *IEEE Transactions on Neural Networks and Learning Systems*, in press
[44] N.N. Karnik and J.M. Mendel, "Introduction to type-2 fuzzy logic systems," in *Proc. IEEE Intl. Conf. on Fuzzy Systems, IEEE World Congress on Computational Intelligence*, 1998, vol. 2, pp. 915–920.
[45] A. Kehagias, "Some remarks on the lattice of fuzzy intervals," *Information Sciences*, vol. 181, no. 10, pp. 1863–1873, 2011.
[46] A. Kehagias and M. Konstantinidou, "L-fuzzy valued inclusion measure, L-fuzzy similarity and L-fuzzy distance," *Fuzzy Sets and Systems*, vol. 136, no. 3, pp. 313–332, 2003.
[47] G.J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ: Prentice Hall, 1995.
[48] L.T. Kóczy and K. Hirota, "Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases," *Information Sciences*, vol. 71, no. 1-2, pp. 169–201, 1993.
[49] L.T. Kóczy and K. Hirota, "Approximate reasoning by linear rule interpolation and general approximation," *International Journal of Approximate Reasoning*, vol. 9, no. 3, pp. 197–225, 1993.
[50] B. Kosko, "Fuzzy entropy and conditioning," *Information Sciences*, vol. 40, no. 2, pp. 165–174, 1986.
[51] B. Li, P.-Y. Liu, R.-X. Hu, S.-S. Mi, and J.-P. Fu, "Fuzzy lattice classifier and its application to bearing fault diagnosis," *Applied Soft Computing*, vol. 12, no. 6, pp. 1708–1719, 2012.
[52] H. Liu, S. Xiong, and C.-a. Wu, "Hyperspherical granular computing classification algorithm based on fuzzy lattices," *Mathematical and Computer Modelling*, vol. 57, no. 3–4, pp. 661–670, 2013.
[53] E.H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1–13, 1975.
[54] J.M. Mendel, "Type-2 fuzzy sets and systems: an overview," *IEEE Comput. Intel. Magazine*, vol. 2, no. 1, pp. 20–29, 2007.
[55] M. Nachtegael, P. Sussner, T. Mélange, and E.E. Kerre, "On the role of complete lattices in mathematical morphology: From tool to uncertainty model," *Information Sciences*, vol. 181, no. 10, pp. 1971–1988, 2011.
[56] S. Nefti, M. Oussalah, and U. Kaymak, "A new fuzzy set merging technique using inclusion-based fuzzy clustering," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 1, pp. 145–161, 2008.
[57] H.T. Nguyen and E.A. Walker, *A First Course in Fuzzy Logic*, 2nd Edition. Chapman and Hall/CRC, 2000.
[58] S.E. Papadakis and V.G. Kaburlasos, "Piecewise-linear approximation of non-linear models based on probabilistically/possibilistically interpreted intervals' numbers (INs)," *Information Sciences*, vol. 180, no. 24, pp. 5060–5076, 2010.
[59] W. Pedrycz, A. Skowron, and V. Kreinovich (eds.), *Handbook of Granular Computing*. Chichester, England: John Wiley & Sons, 2008.
[60] V. Petridis and V.G. Kaburlasos, "Fuzzy lattice neural network (FLNN): a hybrid model for learning," *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 877–890, 1998.
[61] G.X. Ritter and P. Sussner, "An introduction to morphological neural networks," in *Proc. 13th IEEE Intl. Conf. on Pattern Recognition*, 1996, vol. 4, pp. 709–717.
[62] C. Ronse, "Why mathematical morphology needs complete lattices," *Signal Processing*, vol. 21, no. 2, pp. 129–154, 1990.
[63] J. Serra, *Image Analysis and Mathematical Morphology, Vol. 2: Theoretical Advances*. New York, NY: Academic Press, 1988.
[64] B. Šešelja and A. Tepavčević, "Completion of ordered structures by cuts of fuzzy sets: an overview," *Fuzzy Sets and Systems*, vol. 136, no. 1, pp. 1–19, 2003.
[65] B. Šešelja and A. Tepavčević, "Representing ordered structures by fuzzy sets: an overview," *Fuzzy Sets and Systems*, vol. 136, no. 1, pp. 21–39, 2003.
[66] D. Sinha and E.R. Dougherty, "Fuzzification of set inclusion: theory and applications," *Fuzzy Sets and Systems*, vol. 55, no. 1, pp. 15–42, 1993.
[67] P. Sussner and E.L. Esmi, "Morphological perceptrons with competitive learning: Lattice-theoretical framework and constructive learning algorithm," *Information Sciences*, vol. 181, no. 10, pp. 1929–1950, 2011.
[68] P. Sussner and C.R. Medeiros, "An introduction to morphological associative memories in complete lattices and inf-semilattices," in *Proc. of the World Congress on Computational Intelligence (WCCI) 2012, FUZZ-IEEE Program*, Brisbane, Australia, 10-15 June 2012, pp. 1–8.
[69] P. Sussner and M.E. Valle, "Gray-scale morphological associative memories," *IEEE Transactions on Neural Networks*, vol. 17, no. 3, pp. 559–570, 2006.
[70] P. Sussner and M.E. Valle, "Implicative fuzzy associative memories," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, pp. 793–807, 2006.
[71] P. Sussner, M. Nachtegael, T. Mélange, G. Deschrijver, E. Esmi, and E. Kerre, "Interval-valued and intuitionistic fuzzy mathematical morphologies as special cases of L-fuzzy mathematical morphology," *J. Math. Imaging and Vision*, vol. 43, no. 1, pp. 50–71, 2012.
[72] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man & Cybernetics*, vol. 15, no. 1, pp. 116–132, 1985.
[73] M.E. Valle and P. Sussner, "A general framework for fuzzy morphological associative memories," *Fuzzy Sets and Systems*, vol. 159, no. 7, pp. 747–768, 2008.
[74] M.E. Valle and P. Sussner, "Storage and recall capabilities of fuzzy morphological associative memories with adjunction-based learning," *Neural Networks*, vol. 24, no. 1, pp. 75–90, 2011.
[75] M.E. Valle and P. Sussner, "Quantale-based autoassociative memories with an application to the storage of color images," *Pattern Recognition Letters*, in press
[76] M.E. Valle and D.M. Grande Vicente, "Sparsely connected autoassociative lattice memories with an application for the reconstruction of color images," *J. Math. Imaging and Vision*, vol. 44, no. 3, pp. 195–222, 2012.
[77] H. Wu and J.M. Mendel, "Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems," *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 5, pp. 622–639, 2002.
[78] Y. Xu, J. Liu, X. Zhong, and S. Chen, "Lattice-valued matrix game with mixed strategies for intelligent decision suport," *Knowledge-Based Systems*, vol. 32, pp. 56–64, 2012.
[79] Y. Xu, D. Ruan, K. Qin, and J. Liu, *Lattice-Valued Logic*. Heidelberg, Germany: Springer, 2003, ser. Studies in Fuzziness and Soft Computing, vol. 132.
[80] V.R. Young, "Fuzzy subsethood," *Fuzzy Sets and Systems*, vol. 77, no. 3, pp. 371–384, 1996.
[81] L.A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
[82] L.A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 1, pp. 28–44, 1973.

[83] L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning - I," *Information Sciences*, vol. 8, no. 3, pp. 199–249, 1975.
[84] H. Zhang, M. Dong, W. Zhang, and X. Song, "Inclusion measure and similarity measure of intuitionistic and interval-valued fuzzy sets," in *Proc. Intl. Conf. on Intel. Syst. and Knowledge Eng. (ISKE 2007)*, T. Li, Y. Xu, and D. Ruan, Eds. Atlantis Press, 2007.
[85] L. Zhou, "On local inclusion degree of intuitionistic fuzzy sets," in *Rough Sets and Knowledge Technology* (Lecture Notes in Computer Science Series 6954), J. Yao, S. Ramanna, G. Wang, and Z. Suraj, Eds. Berlin, Germany: Springer, 2011, pp. 440–446.