

Novel Analysis and Design of Fuzzy Inference Systems Based on Lattice Theory

Vassilis G. Kaburlasos
Department of Industrial Informatics
Division of Computing Systems
Technological Educational Institution of Kavala
GR-65404 Kavala, Greece
Email: vgkabs@teikav.edu.gr

Athanasios Kehagias
Department of Math., Phys. & Comp. Sciences
Division of Mathematics
Aristotle University of Thessaloniki
GR-54124 Thessaloniki, Greece
Email: kehagiat@hermes.ccf.auth.gr

Abstract—This work presents a Fuzzy Inference System (FIS) as a look-up table for function approximation by interpolation involving Fuzzy Interval Numbers or *FINs* for short. It is shown that the cardinality of the set F of *FINs* equals \aleph_1 , that is the cardinality of the totally ordered lattice R of real numbers. Hence a FIS can implement in principle all $\aleph_2 = 2^{\aleph_1} > \aleph_1$ real functions, moreover a FIS is endowed with a capacity for local generalization. It follows a unification of Mamdani- with Sugeno-type FIS. Based on lattice theory novel interpretations are introduced and, in addition, a tunable metric distance d_k between *FINs* is shown. Several of the proposed advantages are demonstrated experimentally.

I. INTRODUCTION

Even though the notion “fuzzy set” can be defined on any universe of discourse, nevertheless fuzzy sets are typically defined on the real number R universe of discourse, where the name *fuzzy number/interval* is used to denote a convex, normal fuzzy set. For reasons to be explained below, the term *Fuzzy Interval Number* or *FIN* for short is employed here instead of fuzzy number/interval.

Various Fuzzy Inference Systems (FIS) have been developed in practice based mainly either on expert knowledge [13] or on measurements [21]. It turns out that FIS are frequently used in practice for function approximation [23]. In particular a FIS is frequently used for approximating a function $f: R^N \rightarrow R^M$, e.g in automatic control applications [5, 15]. Several publications have compared FIS with various networks for function approximation and learning [7, 12, 14]; note that an account of the latter networks appears in [17].

It is worthwhile noting that the set R of real numbers has emerged from the measurement process. Furthermore note that R is a totally ordered lattice whose cardinality is denoted by \aleph_1 . This work proposes novel perspectives and tools for FIS analysis and design based on a synergy of set theory and mathematical lattice theory.

A critical set-theoretic result here regards the cardinality of *FINs*, where cardinality means how many *FINs* are there. It turns out that there are as many as \aleph_1 *FINs*. It follows that a FIS can implement in principle all $\aleph_2 = 2^{\aleph_1} > \aleph_1$ real

functions; moreover a FIS is endowed with a capacity for local generalization. Based on lattice theory this work introduces a tunable metric distance in the space F of *FINs*, where an integrable *mass function* can be used for tuning. The utility of novel tools is illustrated geometrically on the plane.

The layout of this paper is as follows. Section II summarizes basic Fuzzy Inference Systems (FIS) operation principles. Section III presents useful enhancements in the theory of fuzzy lattices. Section IV deals with Fuzzy Interval Numbers (*FINs*). Section V shows novel perspectives as well as tools for an enhanced FIS analysis and design. Experimental results are demonstrated in section VI. Section VII summarizes the contribution of this work including a discussion of future work.

II. A SUMMARY OF FIS OPERATION PRINCIPLES

A fuzzy inference system (FIS) includes a number of fuzzy rules. For example Fig.1 shows a “Mamdani type” FIS, where the antecedent (IF part) of a rule is the conjunction of N fuzzy statements moreover the consequent (THEN part) of a rule is a single fuzzy statement. A typical input vector $x \in R^N$ may activate in parallel all the rules by a *fuzzification* procedure. The fuzzy consequents of all activated rules are combined and, finally, a single number is produced by a *defuzzification* procedure.

Other types of FIS than a Mamdani type can be obtained for different types of rule consequents. For instance, using an algebraic expression $y = f(x_1, \dots, x_N)$ as a consequent to a rule, a Sugeno type FIS results in [21].

A FIS is frequently used in practice as a device for implementing a function $f: R^N \rightarrow R^M$. The design of a FIS concerns, first, the computation of the parameters which specify both the location and the shape of the fuzzy sets involved in the (fuzzy) rules of a FIS and, second, it may also concern the computation of the parameters of the consequent algebraic equations involved in a Sugeno type FIS. In the aforementioned sense the design of a FIS boils down to an optimal parameter estimation problem, frequently with constraints, moreover a linguistic interpretation is retained [23, 24].

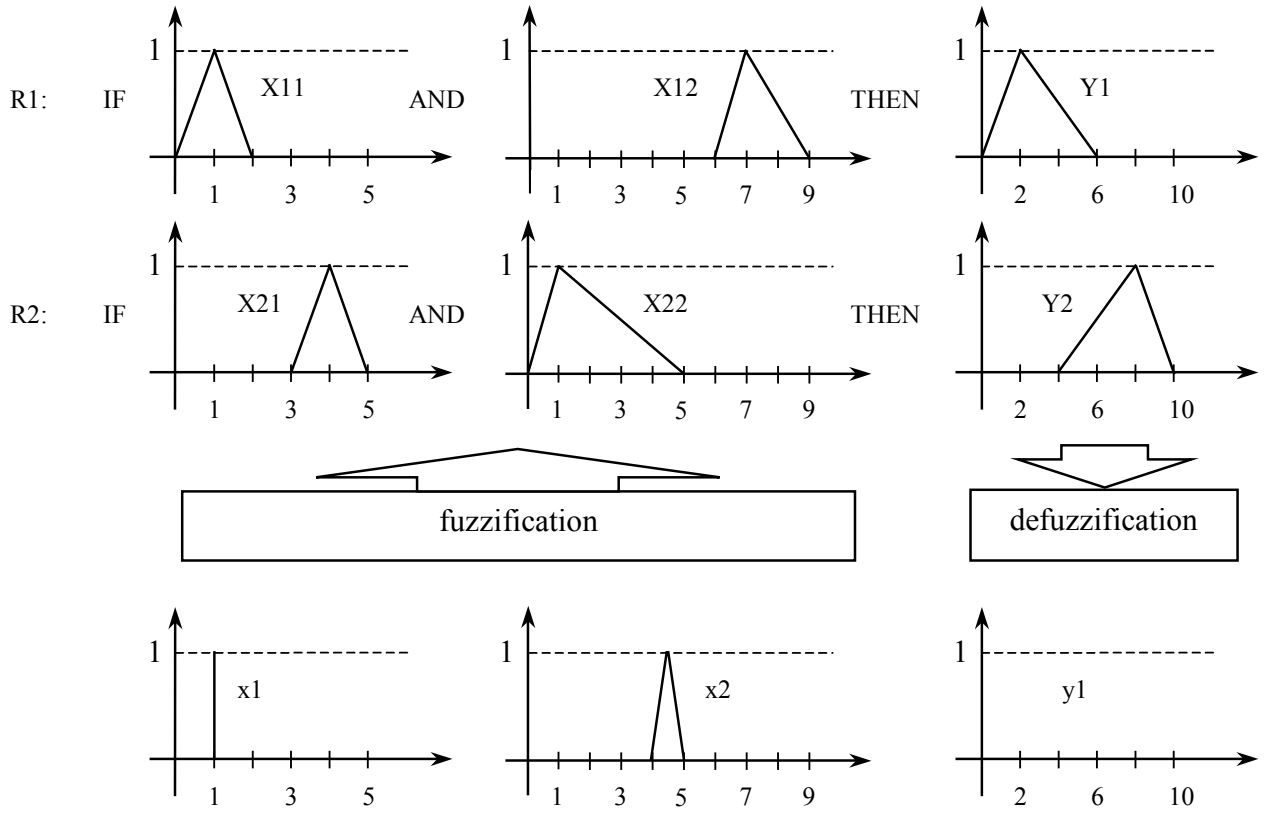


Fig. 1 A Mamdani type FIS with two (fuzzy) rules R1 and R2, two inputs x1 and x2, one output y1. The above FIS, including both a fuzzification and a defuzzification procedure, can be used for implementing a function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$. In the context of this work an input may be a fuzzy set for capturing ambiguity in an input.

III. FUZZY LATTICES REVISITED

The *framework of fuzzy lattices* has been proposed for unifying the treatment of disparate types of data [10]. This section summarizes instrumental lattice-theoretic notions and tools, moreover novel tools are presented.

A *lattice* is a partially ordered set L any two of whose elements have a greatest lower bound or “meet” denoted by $x \wedge y$, and a least upper bound or “join” denoted by $x \vee y$. We say that x and y are *comparable* when either $x \leq y$ or $y \leq x$; otherwise x and y are *incomparable* symbolically $x \parallel y$.

The interest here is in *fuzzy numbers* defined on the real number \mathbb{R} universe of discourse. Note that \mathbb{R} is a *totally ordered* lattice, i.e. for $x, y \in \mathbb{R}$ it is either $x \leq y$ or $y < x$.

The notion fuzzy lattice has been introduced in order to extend the crisp lattice ordering relation (\leq) to all pairs (x, y) in $L \times L$ including incomparable lattice elements [10]. Such an extended relation may be regarded as a fuzzy set on the universe of discourse $L \times L$.

Definition 1: A *fuzzy lattice* is a pair $\langle L, \mu \rangle$, where L is a lattice and μ is a fuzzy relationship $\mu: L \times L \rightarrow [0, 1]$ such that $\mu(x, y) = 1 \Leftrightarrow x \leq y$.

Definition 1 is different from the “standard” definition of a fuzzy lattice first introduced in [1] and later used and generalized by many authors (for a recent review see [22]). It is also different from the approach in [11] regarding the synthesis of fuzzy multivalued connectives. A fuzzy lattice here is defined through an *inclusion measure* function.

Definition 2: Given a lattice L , an *inclusion measure* is a fuzzy relation $\sigma: L \times L \rightarrow [0, 1]$ such that the following conditions are satisfied for every for $u, w, x, y \in L$.

C1. $\sigma(x, x) = 1$.

C2. $x \wedge y < x \Rightarrow \sigma(x, y) < 1$.

C3. $z \leq x \Rightarrow \sigma(y, z) \leq \sigma(y, x)$ - *Consistency Property*

We remark that various “inclusion measures” have been presented in the literature for dealing with a lattice of sets [3, 11, 19]. Our definition here is more general, since it applies not only to fuzzy sets but to elements of a general lattice. In this context $\sigma(x, y)$ can be interpreted as the (fuzzy) degree to which x is less than y ; therefore the notations $\sigma(x, y)$ and $\sigma(x \leq y)$ will be used interchangeably.

The definitions for both a *metric distance* and a *positive valuation* function are shown in the following.

Definition 3: A metric distance in a set \mathcal{S} is a real function $d: \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$ which satisfies: (MD0) $d(x,y) \geq 0$, (MD1) $d(x,y)=0 \Leftrightarrow x=y$, (MD2) $d(x,y) = d(y,x)$, and (MD3) $d(x,y) \leq d(x,z) + d(z,y)$ - *Triangle Inequality*, $x,y,z \in \mathcal{S}$.

Definition 4: A valuation in a lattice L is a function $v: L \rightarrow \mathbb{R}$ such that $v(x)+v(y)=v(x \wedge y)+v(x \vee y)$, $x,y \in L$. A valuation is called *positive* if, for all $x,y \in L$, we have $x < y \Rightarrow v(x) < v(y)$.

If L is a lattice and v is a positive valuation function then function $d(x,y) = v(x \vee y) - v(x \wedge y)$ is a metric distance [2]. Next we define two useful inclusion measures.

Definition 5: Let L be a lattice and v be a positive valuation. Then both the functions $k(x,u) = v(u)/v(x \vee u)$, and $s(x,u) = v(x \wedge u)/v(x)$ are inclusion measures.

Note that the inclusion measure $k(x,u)$ is a typical example of degree of “subsethood” [10, 11].

IV. FUZZY INTERVAL NUMBERS (FINs)

This section summarizes, in an improved notation, useful mathematical tools introduced elsewhere [8, 9, 16]. Novel results are also introduced here.

A. Metric Lattices M^h of Generalized Intervals

Definition 6: A generalized interval of height h is a mapping given by

$$\begin{aligned} 1) & \text{ If } x_1 \leq x_2 \text{ then } \mu_{[x_1, x_2]^h}(x) = \begin{cases} h, & x_1 \leq x \leq x_2 \\ 0, & \text{otherwise} \end{cases}, \text{ else} \\ 2) & \mu_{[x_1, x_2]^h}(x) = \begin{cases} -h, & x_1 \geq x \geq x_2 \\ 0, & \text{otherwise} \end{cases}, \text{ where } h \in (0, 1]. \end{aligned}$$

A generalized interval will be denoted, more compactly, as $[x_1, x_2]^h$. More specifically, if $x_1 \leq x_2$ ($x_1 > x_2$) then $[x_1, x_2]^h$ is called *positive (negative) generalized interval*. The set of positive (negative) generalized intervals of height h will be denoted by M_+^h (M_-^h). The set of *positive (negative)* generalized intervals will be denoted by M_+ (M_-).

The *support* (of a generalized interval) is a function, which maps a generalized interval to its conventional interval support set. An *ordering* relation can be defined in M^h , $h \in (0, 1]$ as shown in the following.

$$\begin{aligned} (OR1) & [a, b]^h \leq [c, d]^h \Leftrightarrow \text{support}([a, b]^h) \subseteq \text{support}([c, d]^h), \\ & \text{for } [a, b]^h, [c, d]^h \in M_+^h \\ (OR2) & [a, b]^h \leq [c, d]^h \Leftrightarrow \text{support}([c, d]^h) \subseteq \text{support}([a, b]^h), \\ & \text{for } [a, b]^h, [c, d]^h \in M_-^h, \text{ and} \\ (OR3) & [a, b]^h \leq [c, d]^h \Leftrightarrow \text{support}([a, b]^h) \cap \text{support}([c, d]^h) \neq \emptyset \\ & \text{for } [a, b]^h \in M_-^h, [c, d]^h \in M_+^h. \end{aligned}$$

It can be shown that the partially ordered set M^h of generalized intervals is a mathematical lattice. More specifically, $[a, b]^h \wedge [c, d]^h = [a \vee c, b \wedge d]^h$ and $[a, b]^h \vee [c, d]^h = [a \wedge c, b \vee d]^h$.

In the totally-ordered lattice \mathbb{R} of real numbers any strictly increasing function is a positive valuation function, the latter can be used for introducing a positive valuation function in M^h .

Proposition 7: Let $f_h: \mathbb{R} \rightarrow \mathbb{R}$ be a strictly increasing function, namely *underlying positive valuation function*. Then function $v: M^h \rightarrow \mathbb{R}$ given by $v([a, b]^h) = f_h(b) - f_h(a)$ is a positive valuation function in M^h .

It follows a novel metric distance $d_h([a, b]^h, [c, d]^h)$ between two generalized intervals $[a, b]^h$ and $[c, d]^h$ given by $d_h([a, b]^h, [c, d]^h) = [f_h(a \vee c) - f_h(a \wedge c)] + [f_h(b \vee d) - f_h(b \wedge d)]$. An underlying positive valuation function $f_h: \mathbb{R} \rightarrow \mathbb{R}$ will be constructed here from an integrable, positive *mass function*

$m_h: \mathbb{R} \rightarrow \mathbb{R}^+$ using formula $f_h(x) = \int_0^x m_h(t) dt$. We point out

that the latter integral is positive (negative) for $x > 0$ ($x < 0$). For example note that mass function $m_h(t) = h$ implies the metric distance $d_h([a, b]^h, [c, d]^h) = h(|a - c| + |b - d|)$ between two generalized intervals $[a, b]^h$ and $[c, d]^h$.

B. The Metric Lattice F of FINs

Definition 8: A *Fuzzy Interval Number (FIN)* is a function either $F: (0, 1] \rightarrow M_+$ (*positive FIN*), or $F: (0, 1] \rightarrow M_-$ (*negative FIN*) such that (1) $F(h) \in M^h$, and (2) $h_1 \leq h_2 \Rightarrow \text{support}(F(h_1)) \supseteq \text{support}(F(h_2))$.

The set of *FINs* is denoted by F ; more specifically the sets of *positive (negative) FINs* will be denoted, respectively, by F_+ (F_-). We remark that a *FIN* is not a fuzzy set, rather a *FIN* is an abstract mathematical notion. There are certain algebraic advantages for negative *FINs*. Positive *FINs* can be interpreted as fuzzy sets. An ordering relation has been introduced in the set F of *FINs* as follows: $F_1 \leq F_2 \Leftrightarrow F_1(h) \leq F_2(h)$, $h \in (0, 1]$. The following proposition introduces a metric in lattice F .

Proposition 9: Let F_1 and F_2 be *FINs* in lattice F . A metric distance function $d_K: F \times F \rightarrow \mathbb{R}$ is given by $d_K(F_1, F_2) = \int_0^1 d_h(F_1(h), F_2(h)) dh$, where $d_h(F_1(h), F_2(h))$ is a metric distance between generalized intervals $F_1(h)$ and $F_2(h)$.

The metric distance $d_K(.,.)$ has several advantages over alternative metric distances defined between fuzzy sets [4, 6]. For instance in the following we compare metric

distance d_K with another well-known metric distance between convex fuzzy sets given by the following

Minkowski metric $d_p(u,v) = \left(\int_0^1 d_H([u]^a, [v]^a) da \right)^{1/p}$ whose

calculation is based on the Hausdorf metric distance d_H between the a -cuts $[u]^a$ and $[v]^a$ of two fuzzy sets u and v [4]. The metrics d_H and d_h produce quite different results in the space \mathbf{R}^1 . More specifically, $d_H([a,b],[c,d]) = \max\{|a-c|, |b-d|\}$, whereas using mass function $m(t)=1$ it follows that $d_h([a,b],[c,d]) = |a-c| + |b-d|$. There are both theoretical and practical advantages for the employment of d_K over d_p . In particular, from a theoretical point of view, there are \aleph_0 different metrics d_p for all different integer values $p=1,2,3,\dots$ whereas there exist $\aleph_1 = 2^{\aleph_0} > \aleph_0$ different metrics d_K . From a practical point it turns out that the capacity to calculate metric d_K based on any strictly increasing function f_h can potentially produce a much larger (finite) number of metrics d_K thus taking full advantage of the existing digital computer memory resources [9].

C. The Cardinality of FINs

Theorem 10: The cardinality $\text{card}(\mathbf{I}_{[a,b]})$ of the set $\mathbf{I}_{[a,b]}$ of non-decreasing functions on the closed interval $[a,b]$ equals $\text{card}(\mathbf{I}_{[a,b]}) = \aleph_1$.

From theorem 10 it follows the major novel theoretical result of this work, that is the cardinality of the set \mathbf{F} of FINs equals \aleph_1 as it will be detailed rigorously elsewhere.

V. ADVANTAGES IN FUNCTION APPROXIMATION

By “function approximation” here is meant an induction of a function $f: \mathbf{R}^N \rightarrow \mathbf{R}^M$ from n pairs $(x_1, y_1), \dots, (x_n, y_n)$ of training data vectors such that both a minimum mean

square error $\text{MSE} = \sqrt{\sum_{i=1}^n \|f(x_i) - y_i\|^2}$ and a useful

capacity for generalization are attained. Various instruments have been employed in the literature for function approximation including polynomials, ARMA models, statistical regression models, multilayer perceptrons, etc. Lately FIS have proliferated in function approximation applications.

A function approximation problem initially includes a “training phase” involving an optimal estimation of a model parameters, the latter can be regarded as an optimal selection of a model in a family of models. For instance, an optimal (in the MSE sense) estimation of the coefficients of an “order n polynomial” can be regarded as the selection of the best among all polynomials of order n . The question now is how many polynomials of order n are there to choose from? It follows that there are $\aleph_1^n = \aleph_1$ such polynomials [20], where \aleph_1 is the cardinality of the set \mathbf{R}

of real numbers. Based on similar arguments as above it can be shown that the cardinality of all ARMA models equals \aleph_1 , furthermore the cardinality of all multilayer perceptrons equals \aleph_1 [8]. The next theoretical question is how many real functions $f: \mathbf{R}^N \rightarrow \mathbf{R}^M$ are there to be approximated? It turns out that there are $\aleph_2 = \aleph_1^{\aleph_1} = 2^{\aleph_1} >$

\aleph_1 real functions [20]. In other words the cardinality of all aforementioned “conventional” models is an order of infinity less than the cardinality of all real functions. However, the aforementioned models retain an important advantage that is their capacity for generalization.

The previous section has shown that the cardinality of the family \mathbf{F} of FINs equals \aleph_1 . It follows that the cardinality of the family of functions $f: \mathbf{R}^N \rightarrow \mathbf{R}^M$ equals \aleph_2 . Each one of the latter functions can be regarded as a Mamdani type FIS. Hence, there is a one-one correspondence between FIS and real functions $f: \mathbf{R}^N \rightarrow \mathbf{R}^M$. Furthermore, a fuzzification/ defuzzification procedure [18] may imply a capacity for local generalization.

The previous analysis did not adhere to FINs of a specific membership function shape. It appears that any family of shapes, e.g. triangular, bell-shaped, etc., would be equally good because every aforementioned family has cardinality \aleph_1 . Furthermore the previous results are retained by Sugeno type FIS because the fuzzy rule consequents in Sugeno type FIS include \aleph_1 algebraic expressions $y = f(x_1, \dots, x_N)$. In addition, there is a practical advantage of FIS in a function approximation application. That is, using the tunable metric distance $d_K(\dots)$ between FINs it is possible to choose among more metric distance functions [9].

Finally note that conventional fuzzification procedures employ exclusively the inclusion measure function $s(\dots)$ as it will be detailed elsewhere.

VI. EXPERIMENTAL RESULTS

Consider Fig.1, where a Mamdani type FIS is shown including two rules R1 and R2. An input pair (x_1, x_2) is presented including, respectively, a number and a fuzzy set.

None of the fuzzy rules in Fig.1 would be activated using fuzzy logic. Nevertheless, using the distance $d_K(\dots)$, it is possible to compute rigorously a degree of activation of a rule. For instance, consider the FINs X22, x2, and X12 copied in Fig.2(a) from Fig.1. We will compute in the following the metric distances $d_K(X22, x2)$ and $d_K(X12, x2)$ using the two different mass functions shown, respectively, in Fig.2(b) and Fig.2(c). On the one hand, the mass function $m_h(t) = h, h \in (0, 1]$ in Fig.2(b) assumes that all the real numbers are equally important. On the other hand, the mass function $m_h(t) = h2e^{-(t-6)/(1+e^{-(t-6)})^2}$ in Fig.2(c) emphasizes the numbers around $t=6$; the corresponding positive valuation function is $f_h(x) = h[(2/(1+e^{-(x-6)}))-1]$, namely *logistic (sigmoid) function*.

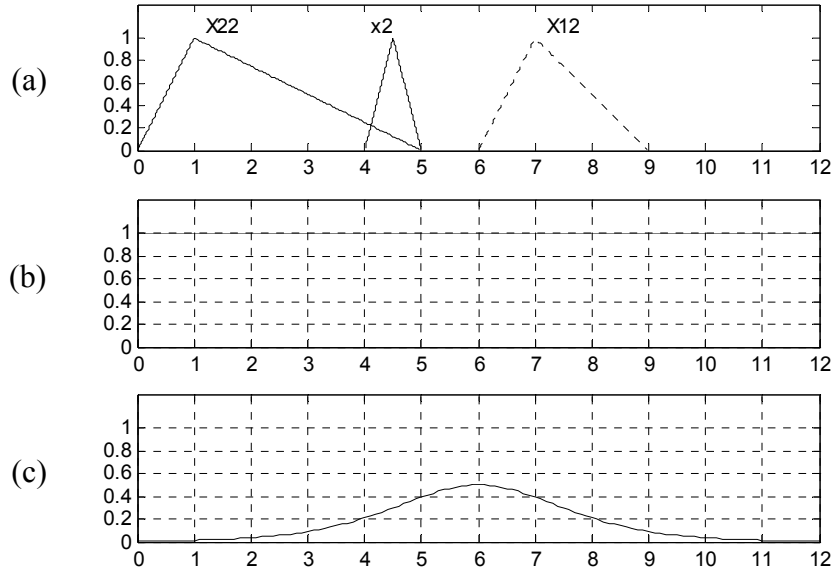


Fig. 2 (a) The *FINs* X12, X22, and x2 have been copied from Fig.1.
 (b) The mass function $m_h(t) = h$, for $h=1$.
 (c) The mass function $m_h(t) = h2e^{-(t-6)^2} / (1 + e^{-(t-6)^2})$, for $h=1$.

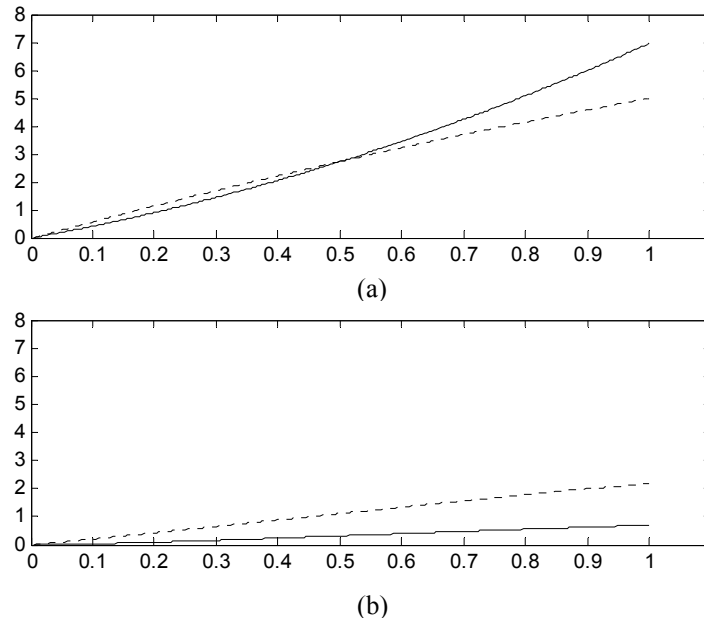


Fig. 3 (a) The metric distance functions $d_K(X22(h), x2(h))$ and $d_K(X12(h), x2(h))$ are plotted in solid and dashed lines, respectively, using the mass function $m_h(t) = h$ shown in Fig.2(b). The area under either curve equals the corresponding distance between two *FINs*. It turns out $d_K(X22, x2) \approx 3.0 > 2.667 \approx d_K(X12, x2)$.
 (b) The metric distance functions $d_K(X22(h), x2(h))$ and $d_K(X12(h), x2(h))$ are plotted in solid and dashed lines, respectively, using the mass function $m_h(t) = h2e^{-(t-6)^2} / (1 + e^{-(t-6)^2})$ shown in Fig.2(c). The area under either curve equals the corresponding distance between two *FINs*. It turns out $d_K(X22, x2) \approx 0.328 < 1.116 \approx d_K(X12, x2)$.

Fig.3(a) plots both functions $d_k(X22(h),x2(h))$ and $d_k(X12(h),x2(h))$ in solid and dashed lines, respectively, using the mass function $m_h(t)=h$. The area under a curve equals the corresponding distance between two *FINs*. It turns out $d_k(X22,x2) \approx 3.0 > 2.667 \approx d_k(X12,x2)$. Fig.3(a) illustrates that for smaller values of h $d_k(X12(h),x2(h))$ is larger than $d_k(X22(h),x2(h))$ and the other way around for larger values of h as expected from Fig.2(a) by inspection.

Fig.3(b) plots both functions $d_k(X22(h),x2(h))$ and $d_k(X12(h),x2(h))$ in solid and dashed lines, respectively, using the mass function $m_h(t)=h2e^{-(t-6)}/(1+e^{-(t-6)})^2$. It turns out $d_k(X22,x2) \approx 0.328 < 1.116 \approx d_k(X12,x2)$.

This example has demonstrated a number of useful tools for tuning FIS design including, first, a mass function $m_h(t)$ can be used for introducing non-linearities in an application. Second, using a metric distance $d_k(.,.)$ it is not necessary to have the whole data domain covered with fuzzy rules. Third, an input to a FIS might be a fuzzy set for dealing with ambiguity in the input data.

VII. DISCUSSION AND CONCLUSION

This work has introduced novel perspectives and tools for Fuzzy Inference System (FIS) analysis and design based on a synergy of set theory and mathematical lattice theory.

A FIS was presented as a look-up table for function approximation by interpolation involving Fuzzy Interval Numbers (*FINs*). The set F of *FINs*, including the fuzzy numbers, was shown to be a metric mathematical lattice. In particular, a tunable metric distance $d_k(.,.)$ was presented in F based on an integrable mass function $m_h(t)$. Furthermore it was shown that the cardinality of the set F of *FINs* equals \aleph_1 , that is the cardinality of the set R of real numbers. Hence a FIS can implement in principle all $\aleph_2 = 2^{\aleph_1} > \aleph_1$

functions $f: R^N \rightarrow R^M$; moreover a FIS is endowed with a capacity for local generalization. Several of the proposed advantages have been demonstrated experimentally including geometric interpretations on the plane.

It has been a common practice in conventional FIS design to optimize the shape and/or the location of the positive *FINs* involved. This work has presented an additional means for tuning the performance of a FIS by a mass function $m_h(t)$ to be employed in future applications.

REFERENCES

- [1] N. Ajmal and K.V. Thomas, "Fuzzy lattices," *Info. Sci.*, vol. 79, pp. 271-291, 1994.
- [2] G. Birkhoff, *Lattice Theory*. Providence, RI: American Math. Society, Colloquium Publications, 25., 1967.
- [3] 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.
- [4] P. Diamond and P. Kloeden, *Metric Spaces of Fuzzy Sets*. World Scientific, Singapore, 1994.
- [5] C. Elkan, "The paradoxical success of fuzzy logic," *IEEE Expert*, vol. 9, no. 4, pp. 3-8, 1994.
- [6] S. Heilpern, "Representation and application of fuzzy numbers," *Fuzzy Sets and Systems*, 91(2):259-268, 1997.
- [7] J.-S. R. Jang and C.-T. Sun, "Functional equivalence between radial basis function networks and fuzzy inference systems," *IEEE Trans. on Neural Networks*, vol. 4, no. 1, pp. 156-159, 1993.
- [8] V.G. Kaburlasos, "Novel fuzzy system modeling for automatic control applications," *Proc. 4th Intl. Conf. on Technology & Automation*, Thessaloniki, Greece, 5-6 October 2002, pp. 268-275.
- [9] V.G. Kaburlasos, "Fuzzy Interval Numbers (*FINs*): Lattice theoretic tools for improving prediction of sugar production from populations of measurements," *IEEE Trans. on Systems, Man and Cybernetics – Part B* (to be published in 2004).
- [10] V.G. Kaburlasos and V. Petridis, "Fuzzy Lattice Neurocomputing (FLN) models," *Neural Networks*, vol. 13, no. 10, pp. 1145-1170, 2000.
- [11] Ath. 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.
- [12] H.M. Kim and J.M. Mendel, "Fuzzy basis functions: comparisons with other basis functions," *IEEE Trans. on Fuzzy Systems*, vol. 3, no. 2, pp. 158-168, 1995.
- [13] E.H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Intl. Journal of Man-Machine Studies*, vol. 7, pp. 1-13, 1975.
- [14] S. Mitaim and B. Kosko, "The shape of fuzzy sets in adaptive function approximation," *IEEE Trans. on Fuzzy Systems*, vol. 9, no. 4, pp. 637-656, 2001.
- [15] K. Passino and S. Yurkovich, *Fuzzy Control*. Reading, MA: Addison Wesley Longman, 1998.
- [16] V. Petridis, and V.G. Kaburlasos, "FLNkNN: Fuzzy Interval Number k-Nearest Neighbor classifier for prediction of sugar production from populations of samples," *Journal of Machine Learning Research*, vol. 4, pp. 17-37, 2003.
- [17] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1481-1497, 1990.
- [18] T.A. Runkler, "Selection of appropriate defuzzification methods using application specific properties," *IEEE Trans. on Fuzzy Systems*, vol. 5, no. 1, pp. 72-79, 1997.
- [19] 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.
- [20] R.R. Stoll, *Set Theory and Logic*. New York, NY: Dover Publications, 1979.
- [21] T. Takagi, and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Systems, Man, and Cybernetics*, vol 15, no. 1, pp. 116-132, 1985.
- [22] A. Tepavcevic and G. Trajkovski, "L-fuzzy lattices: an introduction," *Fuzzy Sets and Systems*, vol. 123, no. 2, pp. 209-216, 2001.
- [23] L.X. Wang and J.M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least-squares learning," *IEEE Trans. on Neural Networks*, vol. 3, no. 5, pp. 807-814, 1992.
- [24] X.-J. Zeng and M.G. Singh, "Knowledge bounded least squares method for the identification of fuzzy systems," *IEEE Trans. on Systems, Man and Cybernetics - Part C*, vol. 33, no. 1, pp. 24-32, 2003.