Universal Approximation Theorem for Uninorm-Based Fuzzy Systems Modeling

Ronald R. Yager [a] Vladik Kreinovich [b]

[a] *Machine Intelligence Institute, Iona College, 715 North Avenue, New Rochelle, NY 10801-1890, USA, email* yager@panix.com

[b] *Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968, USA, email* vladik@cs.utep.edu

Abstract

Most existing universal approximation results for fuzzy systems are based on the assumption that we use t-norms and t-conorms to represent "and" and "or". Yager has proposed to use, within the fuzzy systems modeling paradigm, more general operations based on uninorms. In this paper, we show that the universal approximation property holds for an arbitrary choice of a uninorm.

Key words: Fuzzy connectives and aggregation operations; Fuzzy systems modeling; Universal approximation

1 Introduction

1.1 Uninorms: main idea

One of the main tasks of fuzzy systems modeling is to use expert knowledge formulated in imprecise (fuzzy) terms when modeling systems. An important part of this task is combining multiple pieces of evidence about a statement (hypothesis). This task was recognized early as an important issue in expert system community; see, e.g., (3; 17). Simply stated, the problem is as follows: If two independent experts (or two independent pieces of evidence) support the hypothesis H, A with degree a, and B with degree b, then what is our resulting degree of confidence in H?

If these degrees a and b are the only information we have, then the resulting resulting degree of confidence must depend only on these two values a and b. In mathematical terms, we can say that the resulting degree of confidence

must be a *function* of two real values a and b. In this paper, we will denote the corresponding function by $a * b$. Which function should we choose?

Historically the first approach to this problem came from fuzzy logic. From the logical viewpoint, in our case, H is true if and only if either support from A is true *or* support from B is true. Thus, as the resulting degree of confidence $a * b$ in H, we can take the value $f_\vee(a, b)$, where $f_\vee(x, y)$ is a *t-conorm*, an extension of a 2-valued logical "or" operation to $[0, 1]$-valued fuzzy logic (see, e.g., (8; 14)).

Definition 1. *A function $f_\vee : [0, 1] \times [0, 1] \to [0, 1]$ is called a t-conorm if it satisfies the following four conditions for all x, x', y, y', and z:*

(1) $f_\vee(0, x) = x$;
(2) $f_\vee(x, y) = f_\vee(y, x)$ *(commutativity);*
(3) $f_\vee(x, f_\vee(y, z)) = f_\vee(f_\vee(x, y), z))$ *(associativity);*
(4) if $x \le x'$ *and* $y \le y'$, *then* $f_\vee(x, y) \le f_\vee(x', y')$ *(monotonicity).*

For "and"-type combination, we can use a similar notion of a t-norm:

Definition 2. *A function $f_\& : [0, 1] \times [0, 1] \to [0, 1]$ is called a t-norm if it satisfies the following four conditions for all x, x', y, y', and z:*

(1) $f_\&(1, x) = x$;
(2) $f_\&(x, y) = f_\vee(y, x)$ *(commutativity);*
(3) $f_\&(x, f_\&(y, z)) = f_\&(f_\&(x, y), z))$ *(associativity);*
(4) if $x \le x'$ *and* $y \le y'$, *then* $f_\&(x, y) \le f_\&(x', y')$ *(monotonicity).*

T-conorms and t-norms are not always a completely adequate way of describing the evidence combination. Indeed, according to the properties of a t-conorm, we always have $f_\vee(a, b) \ge a$ and $f_\vee(a, b) \ge b$; as a result, the degree of confidence in H coming from two supporting pieces of evidence is larger (or the same) as the degree of confidence in H coming from just one piece of evidence. This conclusion makes sense if we only allow *supporting* evidence, but sometimes, we encounter evidence which supports the *negation* $\neg H$ to the hypothesis. Intuitively, if we have two difference pieces of negative evidence, then we should *decrease* our degree of the confidence in the hypothesis, i.e., we should have $a * b < a$; on the other hand, if we choose a t-conorm, we get $a * b \ge a$. Thus, for combining pieces of evidence, we need operations which are more general than t-conorms.

First such operations were used already in the historically first expert system MYCIN (see, e.g., (3; 17)). In this system, degrees of confidence take values from the interval $[-1, 1]$ instead of the more traditional $[0, 1]$. Here, negative values describe negative evidence, and positive values describe positive evi-

dence. As we move from -1 to 1, we go from the evidence which absolutely 100% supports the negation $\neg H$ to evidence which slightly supports $\neg H$ to evidence which slightly supports H to evidence which absolutely 100% supports H. The combination operation is defined as follows:

$$x * y = \frac{x + y}{1 + x \cdot y}. \tag{1}$$

For this operation, two pieces of positive evidence increase our degree of confidence ($x * y > x, y$ if $x, y > 0$), while two pieces of negative evidence decrease our degree of confidence ($x * y < x, y$ if $x, y < 0$).

MYCIN's combination operation is defined for all possible pairs (x, y), with one exception of a pair $(-1, 1)$, for which the above formula is not continuously defined. This exception, however, makes perfect sense: the situation when $a = 1$ and $b = -1$ means that we have two pieces of evidence A and B such that A leads to our 100% degree of confidence in H, while B leads to a 100% confidence in $\neg H$, i.e., in a *rejection* of the hypothesis. In this case, we clearly have a contradiction, both degrees cannot be true, so, instead of trying the combine the two inconsistent pieces of evidence, we should try to analyze and correct the inconsistent degrees of confidence a and b.

The MYCIN operation can be easily re-defined on the interval $[0, 1]$ if we map $[-1, 1]$ onto $[0, 1]$ by a mapping $x \to (x+1)/2$. To find $x * y$ for $x, y \in [0, 1]$, we first find the corresponding values $x' = 2x - 1$ and $y' = 2y - 1$ in the interval $[-1, 1]$, then combine x' and y' according to the original MYCIN rule, getting $z' = (x' + y')/(1 + x' \cdot y')$, and then find $z = x * y$ by mapping the resulting value $z' \in [-1, 1]$ back into the interval $[0, 1]$: $z = (z' + 1)/2$. As a result, we get the following operation on the interval $[0, 1]$:

$$x * y = \frac{x \cdot y}{2x \cdot y - x - y} = \frac{x \cdot y}{x \cdot y + (1 - x) \cdot (1 - y)}. \tag{2}$$

More recently, a new approach to describing combination operations has been proposed in (19) under the name of a *uninorm*. The authors of (19) looked into which part of the standard definition of a t-conorm can be weakened in such a way that it allows MYCIN-type operations as well as usual t-conorms.

Among the conditions from Definition 1, the first condition seems to be the most eligible for changing: this condition makes sense if 0 corresponds to the absence of confirmation, but now 0 stands for the largest negative confirmation, so we have to reformulate this condition by using some value $g \in [0, 1]$ which does represent neither positive nor negative confirmation. As a result, we arrive at the following definition:

Definition 3. (1; 2; 7; 18; 19) *A function* $U : [0, 1] \times [0, 1] \to [0, 1]$ *is called a*

uninorm if the following four conditions are satisfied for all x, x', y, y', and z:

- $U(g, x) = x$ *(identity);*
- $U(x, y) = U(y, x)$ *(commutativity);*
- $U(x, U(y, z)) = U(U(x, y), z))$ *(associativity);*
- *if* $x \le x'$ *and* $y \le y'$, *then* $U(x, y) \le U(x', y')$ *(monotonicity).*

We note that if $g = 0$, the uninorm is a t-conorm, and if $g = 1$, the uninorm is a t-norm.

According to our motivation, uninorms are a natural generalization of t-conorms. It is reasonable to consider related generalizations of t-norms (18):

Definition 4. *By a relevancy transformation, we mean a function* $h : [0, 1] \times [0, 1] \to [0, 1]$ *for which the following four conditions are satisfied for all* x, x', y, *and* y'*:*

- $h(1, x) = x$;
- $h(0, x) = g$;
- *if* $y < y'$ *then* $h(x, y) \le h(x, y')$ *(monotonicity);*
- *if* $x < x'$ *then* $h(x, y) \le h(x', y)$ *for* $y \le g$ *and* $h(x, y) \ge h(x', y)$ *for* $y \le g$.

In this paper, we show that fuzzy system models based on uninorms and relevancy transformations can indeed describe arbitrary real-life systems; in more precise terms, we show that such models are universal approximators.

1.2 Fuzzy systems modeling: in brief

Fuzzy systems modeling (see, e.g., (13)) is a methodology that translates the expert's if-then rules of the type

$$\text{if } A_i(x) \text{ then } B_i(y), \quad 1 \le i \le N, \tag{3}$$

in which the properties $A_i(x)$ and $B_i(y)$ are described by using words from natural languages (such as "x is small"), into a *crisp model*, i.e., into a function $f : X \to Y$ describing what control we should apply for a given input $x \in X$. In the important case when y is the control value, the desired crisp model is a control strategy.

This methodology consists of three major steps:

- first, we *formalize* each "linguistic" property $A_i(x)$ or $B_i(y)$ as a *fuzzy set*, i.e., as a function $A_i : X \to [0, 1]$ (or $B_i : Y \to [0, 1]$) which describes, for

each object $x \in X$ (correspondingly, $y \in Y$), to what extent this property holds for this x or y (e.g., to what extent x is small);
- then, we *combine* these fuzzy sets into a *fuzzy relation*, i.e. a function $R(x, y) : X \times Y \to [0, 1]$ which describes, for each input $x \in X$ and for each possible output $y \in Y$, to what extent this particular outputs satisfies the expert's rules;
- finally, we apply some *defuzzification procedure* to the *fuzzy relation* $R(x, y)$, and get the desired control strategy, as a function $\widetilde{f} : X \to Y$.

In most practical application of fuzzy control, *Mamdani's* approach is used in the combination (second) step. In this approach, the fuzzy relation $R(x, y)$ is represented by a logical formula

$$(A_1(x) \,\&\, B_1(y)) \vee \ldots \vee (A_N(x) \,\&\, B_N(y)), \tag{4}$$

where '$\&$' and '$\vee$' stand for connectives of conjunction and disjunction respectively. The actual choice of model requires that we select an interpretation: a t-norm $f_\& : [0, 1] \times [0, 1] \to [0, 1]$ for conjunction and a t-conorm $f_\vee : [0, 1] \times [0, 1] \to [0, 1]$ for disjunction (see, e.g., (8; 14)), and use these operations in the formula (4), resulting in:

$$R(x, y) = f_\vee(f_\&(A_1(x), B_1(y)), \ldots, f_\&(A_N(x), B_N(y))), \tag{5}$$

For every x, we can then apply an appropriate *defuzzification procedure* to $R(x, y)$ and get the desired value y. In this paper, we prove the universal approximation result for such Mamdani-style fuzzy models.

Comment. From the logical viewpoint, it is somewhat more natural to represent the fuzzy relation $R(x, y)$ as a conjunction of implications:

$$(A_1(x) \to B_1(y)) \& \ldots \& (A_N(x) \to B_N(y)), \tag{6}$$

In this case, we select the interpretation: $f_\& : [0, 1] \times [0, 1] \to [0, 1]$ for conjunction and $f_\to : [0, 1] \times [0, 1] \to [0, 1]$, for implication and use these operations, resulting in:

$$R_\to(x, y) = f_\&(f_\to(A_1(x), B_1(y)), \ldots, f_\to(A_N(x), B_N(y))). \tag{7}$$

It is desirable to extend our results to the corresponding fuzzy inference-based fuzzy models.

1.3 Known universal approximation results

In order to guarantee that this methodology can indeed describe an arbitrary system, it is desirable to check that this methodology is *universal*, i.e., that

for any choice of t-norm and t-conorm, for an arbitrary control function $f : X \to Y$, and for an arbitrary accuracy, there exist appropriate if-then rules for which the resulting control strategy represented by $\widetilde{f}(x)$ approximates the original control function $f(x)$ within the given accuracy.

There exists many universal approximation results for fuzzy systems based on t-norms and t-conorms; first such results were formulated and proved, almost simultaneously, in 1990–92 papers by J. Buckley, Z. Cao, E. Czogala, D. Dubois, M. Grabisch, J. Han, Y. Hayashi, C.-C. Jou, A. Kandel, B. Kosko, J. Mendel, H. Prade, and L.-X. Wang; for a recent survey, see, e.g., (10) and references therein.

Comment. There also exist several universal approximation results for implication-style fuzzy models (4; 5; 6; 15; 16).

1.4 Uninorms in fuzzy systems modeling

In (18), the author proposes to use, in fuzzy systems modeling, uninorms and relevancy transformation operations instead of t-conorms and t-norms. Specifically, since the operation U is associative, we can unambiguously define $U(x, y, z, \ldots, t)$ as, e.g., $U(U(U(x, y), z) \ldots, t)$. In (18), the author uses these operators to define

$$R(x, y) = U(h(A_1(x), B_1(y)), \ldots, h(A_N(x), B_N(y))), \tag{8}$$

In this paper, we show that for every pair (h, U), the resulting fuzzy system modeling methodology has the universal approximation property.

2 General Case: Functions Defined on an Arbitrary Compact Set

Let X be a compact metric space with a metric d_X, and let Y be a complete metric space with a metric d_Y. In this section, we will show that uninorm-based systems can approximate an arbitrary continuous function $f : X \to Y$.

Definition 5. *By a defuzzification procedure, we mean a mapping $D : [0, 1]^Y \to Y$ which maps every membership function $\mu : Y \to [0, 1]$ (which is not identically zero) into an element $D(\mu) \in Y$ for which $\mu(D(\mu)) > 0$.*

Theorem 1. *Let h be a relevancy transformation, and let U be a uninorm. Then, for every compact metric space X, for every continuous function f :*

$X \to Y$ into a complete metric space Y, and for every real number $\varepsilon > 0$, there exist fuzzy rules of type (3) *for which, for each defuzzification procedure D, the result $\widetilde{f}(x)$ of defuzzifying the relation R (obtained using h and U) is ε-close to f.*

Comments.

- For the convenience of the readers, all the proofs are placed in the special Proofs section.
- Since uninorms are more general than t-norms and t-conorms, the universal approximation theorems for t-norms and t-conorms are a particular case of this result. Theorem 1 extends these two results to a more general case of uninorms.

3 Case When $Y = \mathbb{R}$

In the case when Y is a real line ($Y = \mathbb{R}$), we can use a different class of possible "defuzzification procedures" and still get the same universal approximation result. Namely, we can use the following definition:

Definition 5′. ($Y = \mathbb{R}$) *By a defuzzification procedure, we mean a mapping D which maps every non-zero membership function $\mu : \mathbb{R} \to [0, 1]$ into an real number $D(\mu)$ in such a way that for an arbitrary interval $[a, b]$, if a membership function $\mu(x)$ is equal to 0 for all values x outside an interval $[a, b]$, then $D(\mu) \in [a, b]$.*

Comments.

- Both centroid and center-of-maximum are defuzzification procedures in this sense.
- The main difference between Definition 5 and Definition 5′ is that in Definition 5′, the interval $[a, b]$ might still contain x for which $\mu(x) = 0$. This is true, e.g., for centroid defuzzification, when we apply it to a symmetric membership function $\mu(-x) = \mu(x)$ for which $\mu(0) = 0$ (see, e.g., (20; 21)).

Theorem 1′. *Let h be a relevancy transformation, and let U be a uninorm. Then, for every compact metric space X, for every continuous function $f : X \to \mathbb{R}$ and for every real number $\varepsilon > 0$, there exist fuzzy rules of type* (3) *for which, for each defuzzification procedure D, the result $\widetilde{f}(x)$ of defuzzifying the relation R (obtained using h and U) is ε-close to f.*

4 Proofs

4.1 Proof of Theorem 1

4.1.1 General structure of the proof

This proof is similar to the original Kosko's proof (9) of a universal approximation result for a t-norm and t-conorm, and to our own proofs from (11; 12).

This proof uses the following auxiliary definition:

Definition 6. *Let $f : X \to Y$ be a continuous function from X to Y, and $\varepsilon > 0$ be a real number. We say that a fuzzy relation $R : X \times Y \to [0,1]$ ε-approximates a function $f : X \to Y$ if the following two conditions hold:*

- *for every $x \in X$, $R(x, f(x)) > 0$, and*
- *for every $x \in X$ and $y \in Y$, if $R(x,y) > 0$, then $d_Y(y, f(x)) \le \varepsilon$.*

Then, Theorem 1 follows from the following two propositions:

Proposition 1. *Let h be a relevancy transformation, and let U be a uninorm. Then, for every compact metric space X, for every continuous function $f : X \to Y$ into a complete metric space Y, and for every real number $\varepsilon > 0$, there exist fuzzy rules of type* (3) *for which the fuzzy relations R (obtained using h and U) ε-approximate f.*

Proposition 2. *If a fuzzy relation $R(x,y)$ ε-approximates a function $f(x)$, then, for every defuzzification procedure D, the result $\widetilde{f}(x) = D(\mu_x)$ of applying this defuzzification procedure D to the corresponding membership function $\mu_x(y) = R(x,y)$ is ε-close to $f(x)$, i.e., $d_Y\left(\widetilde{f}(x), f(x)\right) \le \varepsilon$.*

4.1.2 Proof of Proposition 1

1°. Let us first select the rules for which, as we will show, the corresponding relation (8) approximates a given function f.

Let us take $\varepsilon_1 = \varepsilon/2$. Since f is continuous on a compact set X, it is also uniformly continuous. Therefore, there exists $\delta > 0$ such that if $d_X(x,x') \le \delta$, then $d_Y(f(x), f(x')) \le \varepsilon_1$.

Since X is a compact metric space, there exists a finite δ-net for X, i.e., a

finite set of elements $x^{(1)}, \ldots, x^{(N)} \in X$ for which, for every $x \in X$, there exists an i for which $d_X\left(x, x^{(i)}\right) \le \delta$. For each of these elements $x^{(i)}$, we can find $y^{(i)} = f\left(x^{(i)}\right)$. We will show that Proposition 1 holds for N rules of type (3) where for every i,

- $A_i(x) = 1$ if $d_X\left(x, x^{(i)}\right) \le \delta$, and $A_i(x) = 0$ otherwise;
- $B_i(y) = 1$ if $d_Y\left(y, y^{(i)}\right) \le \varepsilon_1$, and $B_i(y) = 0$ otherwise.

Informally, the i-th rule can be described as "if $x \approx x^{(i)}$ then $y \approx y^{(i)}$".

2°. Let us now describe the possible values of the degrees $h(A_i(x), B_i(y))$ with which different rules are satisfied.

All these fuzzy sets A_i and B_i are crisp: indeed, $A_i(x)$ is a characteristic function of the inequality $d_X\left(x, x^{(i)}\right) \le \delta$, and $B_i(y)$ is a characteristic function of the inequality $d_Y\left(y, y^{(i)}\right) \le \varepsilon_1$. Thus, each of the functions $A_i(x)$ and $B_i(y)$ can take only values 0 or 1.

For these values, due to the properties of the relevance transformation, $h(1,1) = 1$, $h(1,0) = 0$, and $h(0,0) = h(0,1) = g$. So:

- if $x \approx x^{(i)}$ and $y \approx y^{(i)}$, then $h(A_i(x), B_i(y)) = 1$;
- if $x \approx x^{(i)}$ and $y \not\approx y^{(i)}$, then $h(A_i(x), B_i(y)) = 0$;
- if $x \not\approx x^{(i)}$, then $h(A_i(x), B_i(y)) = g$.

3°. Let us now show that the relation R ε-approximates the given function f.

3.1°. In accordance with the definition of ε-approximation, we first prove that for every $x \in X$, we have $R(x, f(x)) > 0$.

Indeed, let x be an arbitrary element of the set X.

For every i for which $x \not\approx x^{(i)}$, we have $h(A_i(x), B_i(y)) = g$.

For every i for which $x \approx x^{(i)}$, i.e., for which $d_X\left(x, x^{(i)}\right) \le \delta$, by our choice of δ, we have $d_Y\left(f(x), y^{(i)}\right) \le \varepsilon_1$, i.e., $y \approx y^{(i)}$. Thus, for such i, we have $h(A_i(x), B_i(y)) = 1$.

Since $x^{(1)}, \ldots, x^{(N)}$ is a δ-net, there exists an i for which $d_X\left(x, x^{(i)}\right) \le \delta$. Thus, $R(x, y)$ is equal to the uninorm combination of several values, at least one of which is equal to 1, and the rest are equal to either 1 or to g: $R(x, f(x)) = U(1, \ldots, 1, g, \ldots, g)$.

By definition of a uninorm, g acts as an identity ($U(g,x) = x$ for all x), so we can simply delete g's in the uninorm combination: $R(x, f(x)) = U(1,\ldots,1,g,\ldots,g) = U(1,\ldots,1)$.

Since $U(1,g) = 1$ and $g \le 1$, from monotonicity, we conclude that $1 = U(1,g) \le U(1,1)$ hence $U(1,1) = 1$. Thus, $R(x, f(x)) = U(1,\ldots,1) = 1$ hence $R(x, f(x)) > 0$.

3.2°. Let us now prove that for every $x \in X$ and $y \in Y$, if $R(x,y) > 0$, then $d_Y(y, f(x)) \le \varepsilon$.

Indeed, we already know that $R(x,y)$ is a uninorm combination of several values 0, 1, and g. We also know that since $x \approx x^{(i)}$ for some i, at least one of these values is either 0 or 1. We can ignore g's in this combination, so $R(x,y)$ is a uninorm combination of several 0's and 1's.

For a uninorm, $U(g,0) = 0$ and so, due to monotonicity, $U(0,0) \le U(g,0) = 0$ and $U(0,0) = 0$. Thus, if all of the combined values were 0's, we would have $R(x,y) = 0$. Since we have $R(x,y) > 0$, this means that at least one of the combined values is equal to 1. This means that there exists an i for which $d_X\left(x, x^{(i)}\right) \le \delta$ and $d_Y\left(y, y^{(i)}\right) \le \varepsilon_1$. Due to our choice of δ, from $d_X\left(x, x^{(i)}\right) \le \delta$, we can conclude that $d_Y\left(f(x), f\left(x^{(i)}\right)\right) = d_Y\left(f(x), y^{(i)}\right) \le \varepsilon_1$. Thus, from the triangle inequality, we conclude that $d_Y(y, f(x)) \le d_Y\left(y, y^{(i)}\right) + d_Y\left(y^{(i)}, f(x)\right) \le \varepsilon_1 + \varepsilon_1 = \varepsilon$. The statement is proven, and so is Proposition 1.

4.1.3 Proof of Proposition 2

By definition, $\widetilde{f}(x) = D(\mu_x)$, where the membership function $\mu_x : Y \to [0,1]$ is defined as $\mu_x(y) = R(x,y)$. By the definition of a defuzzification procedure, for every $x \in X$, we have $\mu_x(D(\mu_x)) > 0$, i.e., by definition of μ_x, $R\left(x, \widetilde{f}(x)\right) > 0$. From the definition of ε-approximation, we can now conclude that $d_Y\left(\widetilde{f}(x), f(x)\right) \le \varepsilon$. The proposition is proven, and so is Theorem 1.

4.2 Proof of Theorem 1′

For $Y = \mathbb{R}$, Proposition 1 is still true. So, the only thing that needs to be proven anew is the new version of Proposition 2 (corresponding to the new definition of defuzzification).

Indeed, by definition, $\widetilde{f}(x) = D(\mu_x)$, where the membership function $\mu_x : Y \to$

$[0,1]$ is defined as $\mu_x(y) = R(x,y)$. From the definition of ε-approximation, we conclude that if $R(x,y) > 0$, then $d_Y(y, f(x)) = |y - f(x)| \leq \varepsilon$. Thus, if $|y - f(x)| > \varepsilon$, we have $R(x,y) = \mu_x(y) = 0$. Hence, the function $\mu_x(y)$ is equal to 0 outside the interval $[f(x) - \varepsilon, f(x) + \varepsilon]$. By the new definition of a defuzzification procedure, we can now conclude that the result $\widetilde{f}(x)$ of its defuzzification also belongs to the same interval, i.e., that $\left|\widetilde{f}(x) - f(x)\right| \leq \varepsilon$. The proposition is proven, and so is Theorem $1'$.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209 and grant NCC2-1232, NSF grants No. DUE-9750858, CDA-9522207, EAR-0112968, EAR-0225670, and 9710940 Mexico/Conacyt, by the United Space Alliance, grant No. NAS 9-20000 (PWO C0C67713A6), by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grants F49620-95-1-0518 and F49620-00-1-0365, and by the National Security Agency under Grants No. MDA904-98-1-0561.

The authors are very thankful to the anonymous referees for valuable suggestions.

References

[1] B. De Baets, "Uninorms; the known classes", In: D. Ruan, H. A. Abderrahim, P. D'hondt, and E. E. Kerre (eds.), *Fuzzy Logic and Intelligent Technologies for Nuclear Science and Industry*, World Scientific, Singapore, 1998, pp. 21–28.

[2] B. De Baets and J. Fodor, "Residual operators of uninorms", *Soft Computing*, 1999, Vol. 3, pp. 89–100.

[3] B. G. Buchanan and E. H. Shortliffe, *Rule-Based Expert Systems*, Addison-Wesley, Reading, MA, Menlo Park, CA, 1984.

[4] D. Dubois, M. Grabisch, and H. Prade, "Gradual rules and the approximation of functions", *Proceedings of the 2nd International Conference on Fuzzy Logic and Neural Networks*, Iizuka, Japan, July 17–22, 1992, pp. 629–632.

[5] D. Dubois, M. Grabisch, and H. Prade, "Synthesis of real-valued mappings based on gradual rules and interpolative reasoning", *Proc. of the IJCAI'93 Fuzzy Logic in AI Workshop*, Chaméry, France, Aug. 28–Sept. 3, 1993, pp. 29–40.

[6] D. Dubois, H. Prade, and M. Grabisch, "Gradual rules and the approximation of control laws", In: H. T. Nguyen, M. Sugeno, R. Tong, and R. Yager (eds.), *Theoretical aspects of fuzzy control*, J. Wiley, N.Y., 1995, pp. 147–181.

[7] J. Fodor, R. R. Yager, and A. Rybalov, "Structure of uni-norms", *Internat. J. Uncertainty, Fuzziness, and Knowledge-Based Systems*, 1997, Vol. 5, pp. 411–427.

[8] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, NJ, 1995.

[9] B. Kosko, "Fuzzy systems as universal approximators", *Proceedings of the IEEE International Conference on Fuzzy Systems 1992*, San Diego, CA, 1992, pp. 1153–1162.

[10] V. Kreinovich, G. C. Mouzouris, and H. T. Nguyen, "Fuzzy rule based modeling as a universal approximation tool", In: H. T. Nguyen and M. Sugeno (eds.), *Fuzzy Systems: Modeling and Control*, Kluwer, Boston, MA, 1998, pp. 135–195.

[11] H. T. Nguyen and V. Kreinovich, "On approximations of controls by fuzzy systems", *Fifth International Fuzzy Systems Association World Congress*, Seoul, Korea, July 1993, p. 1414–1417.

[12] H. T. Nguyen, V. Kreinovich, and O. Sirisaengtaksin, "Fuzzy control as a universal control tool", *Fuzzy Sets and Systems*, 1996, Vol. 80, No. 1, pp. 71–86.

[13] H. T. Nguyen and M. Sugeno (eds.), *Fuzzy Systems: Modeling and Control*, Kluwer, Boston, MA, 1998.

[14] H. T. Nguyen and E. A. Walker, *First Course in Fuzzy Logic*, CRC Press, Boca Raton, FL, 1999.

[15] V. Novák and I. Perfilieva, "Evaluating Linguistic Expressions and Functional Fuzzy Theories in Fuzzy Logic", In: L. A. Zadeh and J. Kacprzyk (eds.), *Computing with Words in Information/Intelligent Systems*, Springer-Verlag, Berlin, 1999, pp. 383–406.

[16] I. Perfilieva, "Normal Forms for Fuzzy Logic Relations and the Best Approximation Property", *Proc. Conf. EUSFLAT'99*, Palma de Mallorca 1999.

[17] E. H. Shortliffe, *Computer-based medical consultation: MYCIN*, Elsevier, New York, 1976.

[18] R. R. Yager, "Uninorms in fuzzy systems modeling", *Fuzzy Sets and Systems*, 2001, Vol. 122, No. 1, pp. 167–175.

[19] R. R. Yager and A. Rybalov, "Uninorm aggregation operators", *Fuzzy Sets and Systems*, 1996, Vol. 80, pp. 111-120.

[20] J. Yen and N. Pfluger, "Path planning and execution using fuzzy logic", In: *AIAA Guidance, Navigation and Control Conference*, New Orleans, LA, 1991, Vol. 3, pp. 1691–1698.

[21] J. Yen, N. Pfluger, and R. Langari, "A defuzzification strategy for a fuzzy logic controller employing prohibitive information in command formulation", *Proceedings of IEEE International Conference on Fuzzy Systems*,

San Diego, CA, March 1992, pp. 717–723.