Appl. Math. and Comp. Sci., 1996, vol.6, No.3, 505–527

ALGEBRAIC OPERATORS: AN ALTERNATIVE APPROACH TO FUZZY SETS

WŁADYSŁAW HOMENDA*,**

The paper deals with processing of uncertain information in terms of its repetitiveness and negativity. Neither of these features is handled in the lattice structure of fuzzy sets. In the case of repetitive information, the max-min operators are *noninteractive* (see (Pedrycz, 1993)). Triangular norms presented there can cope with repetitive information. Nevertheless, negativity still creates problems. In this paper, a new approach to processing of uncertain information is proposed. This approach is based on some extension of fuzzy sets theory: the linear space and ring structure of fuzzy sets. This extension is created by new operators, algebraic rather than logic. In this paper, the operators are called a-p-v operators. Basic properties of the introduced spaces are discussed. Both repetitive and negative kinds of information are handled in algebraic structures of fuzzy sets. Operators introduced in this paper are applied in two examples: fuzzy reasoning and fuzzy neural networks.

1. Introduction

The basis of the theory of fuzzy sets is commonly related to classical set theory. Namely, given the universe of discourse X one can define a fuzzy set A over X as a mapping $\mu_A : X \to [0,1]$ and interpret the values of this mapping for each element of X as the grades of membership of fuzzy set A. Apparently, this is an extension of classical set theory, which gives the respective interpretation of a non-fuzzy set A in the universe of discourse X through the characteristic function $\chi_A : X \to \{0,1\}$.

Following this intuitive analogy, we get the space of fuzzy sets

$$F(X) = \Big\{\mu : X \to [0,1]\Big\}$$

which is equipped with operators similar to those of set theory: the union, intersection and complement. To be more specific, the tuplet

$$\Big(F(X), \cup, \cap, ^{-}\Big)$$

* Centro de Investigacion Cientifica y de Education Superior de Ensenada, Km. 107 Carretera Tijuana/Ensenada, Ensenada, B.C., Mexico 22860, e-mail: homenda@cicese.mx

** On leave from Institute of Mathematics, Warsaw University of Technology, pl. Politechniki 1, 00-661, Warsaw, Poland, e-mail:homenda@alpha.im.pw.edu.pl

with $A \cup B$ defined as the max operation on respective values of membership functions of fuzzy sets A and B, $A \cup B$—as the min operation, and the complement as subtraction from 1, constitute the structure *very similar* to that obtained from the classical set theory. The tuplet $(F(X), \cup, \cap, ^{-})$ is a complete lattice while $(P(X), \cup, \cap, ^{-})$ is a Boolean algebra, where $P(X)$ stands for the family of all crisp sets defined in the universe of discourse X.

A lot of research has been done following the theoretical foundations of the lattice structure of fuzzy sets, including studies on fuzzy control (where most real-life implementations refer to max and min operators), fuzzy reasoning, relational equations, pattern recognition, neural networks, etc. The properties of the max-min based fuzzy set structure are well understood and widely utilized in practice. Nevertheless, there is a disadvantage called by Pedrycz (1993) *noninteractivity* of max and min operations. To avoid this drawback, triangular norms were utilized. Let us follow Pedrycz (1993) to discuss more clearly the problem of noninteractivity of max and min operations:

> "For any arguments a and b from the unit interval $[0, 1]$, the result of $\min(a, b)$ yields a and does not depend on b as far as $b > a$. In other words, in this case the computed value of $\min(a, b)$ is not influenced by the second argument b. While this type of insensitivity could be viewed advantageous in some situations, this phenomenon in general might lead to a fairly undesired performance of the operation. Let us put $a = 0.1$; then for all the values of b varying over broad sub-range of the unit interval, $0.1 < b \leq 1$, its changes do not affect the final result. In the models of the logical connectives for fuzzy sets it has been experimentally verified, cf. (Zimmermann and Zysno, 1980), that some sort of *compensation* (i.e. interaction) between the values of the membership functions does occur. This behaviour clarifies why t-norms other than the minimum operation can be successfully utilized to model interaction between items under aggregation. Secondly, the lattice operations cannot cope with the *repetitive information* (which in fact has a well legitimized statistical meaning)."

To supplement the above excerpt, let us consider the following observation. In the decision making process we often are faced with data coming from sources of different character, e.g. some information is available from experts while other comes from pooling. These sources of information may be regarded as the sources of *repetitive information*. To get an overall picture of all sources of information we have to refer to their aggregation. In this case noninteractivity of aggregation will distort the final result.

Triangular norms, the generalized max-min operators, cf. (Gupta and Rao, 1994; Pedrycz, 1993), still have a drawback related to the nature of the lattice structure—they cannot cope with *negative information*.

Let us explain the meaning of *negative information*. Notice that the degree of membership function, say $A(x_0)$, specifies the grade to which x_0 satisfies the concept conveyed by the fuzzy set A. Observe that $A(x_0)$ set to 1 denotes complete membership, while $A(x_0) = 0$ excludes its membership. Note, however, that nothing is said explicitly about the degree to which an element x_0 does not belong to A. By default it could be stated that as far as $A(x_0)$ decreases, the grade of belongingness of x_0 to the complement of A should increase. This leads to the interpretation that membership values from the interval (0, 0.5) should be considered as the grade of exclusion from the set A: the lower the membership value, the greater the grade of exclusion. This interpretation satisfies the default but is not well-handled by neither simple max-min operators nor by triangular norms. The main drawback of the *lattice-oriented* triangular norms is that they cannot handle negative information in a way intuitively well legitimized: negative information, if present among arguments of the operator, decreases the result of the operation. So, if many pieces of information are aggregated and most of them are *negative* pieces of information, the result is expected to be *negative.* And vice versa, if most of aggregated pieces of information are *positive*, the result is expected to be *positive.*

A simple example related to a reasoning rule underlines and explains the existence of this type of problems.

Example 1. The description of a reasoning rule given by an expert is often of the form:

If for object O features $(F_1, \dots, F_n)$ are observed with grades close to 1,
then object O belongs to class C with grade close to 1

or, more precisely,

If for object O features $(F_1, \dots, F_n)$ are observed with grades $(f_1, \dots, f_n)$,
respectively, then object O belongs to class C with grade equal to c.

Let us make some more assumptions to give evidence of this problem:

- the grades in the reasoning rule $(f_1, \dots, f_n)$ are equal to $(0.9, 0.9, \dots, 0.9)$, respectively;
- the membership grade c is equal to 0.9;
- the features $(F_1, \dots, F_n)$ of the object O were observed with certainty factors $(0.8, 0.2, 0.8, \dots, 0.8)$.

Utilizing the following formula:

$$\min\Big(\max\big(\min(f_1, x_1), \dots, \min(f_n, x_n)\big), c\Big)$$

we obtain 0.8 as the grade of membership of the object O to the class C. The result is acceptable since most of the grades of features *are close to 1*, so the final result reflects significantly the situation. On the other hand, if one deals with collected values of grades of features equal to $(0.2, 0.8, 0.2, \dots, 0.2)$, then the above rule yields again the value 0.8, which, however, according to our expectations, is much too high. But this result is not unexpected because lattice operators applied in the rule cannot cope with this type of problems, i.e. the problems related to *negativity*: most of the grades of observed features are *close to 0*, i.e. the test for their existence gave a *negative* result and, because of this, the expected result should also be negative, *close to 0*.

For simplicity, only max and min operators were used in above rules, but since the following inequalities hold:

$$s(a,b) \geq \max(a,b) \geq \min(a,b) \geq t(a,b)$$

then it is clear that the results obtained by replacing the min operator by any t-norm and the max operator by a t-conorm (denoted here by s) will be still far away from our intuitive expectation. In this formula, the *If a then b* rule is implemented as the $\min(a,b)$ operator. It is easy to check that using other operators for implementing the *If a, then b* rule, (e.g. the Łukasiewicz, Kleene-Dienes or other rules) will give a similar (i.e. *close to 1*) result.

In this paper, we will discuss a new kind of operators in the space of fuzzy sets and present how the above shortcoming can be eliminated.

2. Algebraic Structures of Fuzzy Sets

Membership functions are usually defined as mappings from any universe of discourse into the unit interval. This fact comes from a tradition rather than is a result of theoretical requirements. Membership functions may refer to values from $[-1,1]$. This symmetrical bipolar representation was used in dealing with uncertainty in the MYCIN system (Dubois and Prade, 1980), it is also applied in fuzzy neural networks (Gupta and Rao, 1994), in cognitive maps (Kosko, 1986), etc. Representation of negative information is also more obvious in the case of bipolarity: a positive value defines the grade of inclusion of the element into a fuzzy set, while negative—the grade of exclusion. The greater the absolute value of membership, the more certain inclusion/exclusion information. By treating the intersection and union as implemented by the minimum and maximum, respectively, and viewing the complement as changing the sign of the value of the membership function, the spaces of grades of membership are isomorphic for both cases: unipolar and bipolar. Take for instance the following mappings:

- for crisp sets

$$i:\{0,1\}^X \to \{-1,1\}^X \qquad \text{such that} \qquad i(x)=2x-1,$$

- for fuzzy sets

$$i:[0,1]^X \to [-1,1]^X \qquad \text{and as before} \qquad i(x)=2x-1.$$

These mappings are isomorphic functions, where a straightforward proof of satisfaction of this property can easily be derived, cf. (Homenda and Pedrycz, 1991). The difference between both representations of fuzzy sets, though insignificant from the theoretical point of view, may play a significant role in practice, but this topic is out of the scope of this paper.

Despite the above observations, we will also refer to unipolar representation of fuzzy sets. Bipolar representation of fuzzy sets, though closer to the nature of a *negative* kind of fuzzy information, still seems to be less appropriate than the widely spread unipolar representation.

In the sequel, we adopt the notation

$$F(X) = (0,1)^X = \left\{\mu |\, \mu : X \to (0,1)\right\}$$

$$\underline{F(X)} = [0,1]^X = \left\{\mu |\, \mu : X \to [0,1]\right\}$$

$$G(X) = (-1,1)^X = \left\{\mu |\, \mu : X \to (-1,1)\right\}$$

$$\underline{G(X)} = [-1,1]^X = \left\{\mu |\, \mu : X \to [-1,1]\right\}$$

In the space $\underline{G(X)}$ of membership functions the following observations can be made:

- a fuzzy set is empty if its membership function is identically equal to -1 on X,
- the complement of a fuzzy set A, denoted by A', is defined by $f_{A'} = -f_A$,
- definitions of equality, containment, union and intersection are taken from classical fuzzy set theory,
- the value 1 denotes complete inclusion while the value -1 refers to complete exclusion,
- the greater the absolute value of the membership function, the more certain information about the object; 0 refers to the case where the element is neither accepted to the category nor excluded from it. The higher the absolute value of the grade of membership, the less uncertainty allocated to the corresponding element of the universe of discourse.

Let us consider the following example:

Example 2. Let X be the real line $\mathbb{R}^1$ and let A be a fuzzy set of numbers that are much greater than 1. Then one can give a precise, albeit subjective characterization of A by specifying $f_A(x)$ as a function on $\mathbb{R}^1$. Representative values of such a function might be as shown in Table 1.

Tab. 1. Classical processing of fuzzy sets.

x	f_A	g_A	$f_{A'}$	$g_{A'}$	$f_A \cup f_{A'}$	$g_A \cup g_{A'}$	$f_A \cap f_{A'}$	$g_A \cap g_{A'}$
1	0	−1	1	1	1	1	0	−1
5	0.05	−0.9	0.95	0.9	0.95	0.9	0.05	−0.9
10	0.2	−0.6	0.8	0.6	0.8	0.6	0.2	−0.6
15	0.5	0	0.5	0	0.5	0	0.5	0
30	0.7	0.4	0.3	−0.4	0.7	0.4	0.3	−0.4
100	0.95	0.9	0.05	−0.9	0.95	0.9	0.05	−0.9
500	1	1	0	−1	1	1	0	−1

When the mapping into $[-1, 1]$ is applied, the membership function is denoted by g_A. The values of membership functions f_A and g_A, complements $f_{A'}$ and $g_{A'}$, unions $f_A \cup f_{A'}$ and $g_A \cup g_{A'}$, together with intersections $f_A \cap f_{A'}$ and $g_A \cap g_{A'}$ are outlined in Table 1.

The value $g_A(5) = -0.9$ gives less uncertainty than the value $g_A(30) = 0.4$ as to the exclusion/ inclusion of the elements 5 and 30 to the set A. The element 5 is more certainly excluded from A than the element 30 is included in A. Because of this, the space $\underline{G(X)}$ seems to be more suitable to express the grades of inclusion/exclusion elements from a fuzzy set/concept. Nevertheless, as is noted above, the classical space of fuzzy sets $\underline{F(X)}$ will also be referred to.

On the basis of the above observations, new operators are presented in the subsequent sections. These operators are called *a-p-v* operators (sum-product-vector operators). They create algebraic structures of fuzzy sets: the ring and vector (linear) space.

2.1. Vector Space of Fuzzy Sets

The vector (linear) space of fuzzy sets will be defined by introducing new operators in the space of fuzzy sets, i.e. in the space $G(X)$ (see (Homenda and Pedrycz, 1991) for a primary discussion of this topic). In general, let us introduce a transforming function h from the open interval $(-1, 1)$ onto the real line $\mathbb{R}$, such that it is

- continuous,
- symmetric, i.e. $h(-x) = -h(x)$,
- strictly increasing.

Obviously, we have

- $h(0) = 0$,
- $\lim h(x) = \pm\infty$ as $x \to \pm 1$,
- h^{-1} exists.

2.3. Processing of Fuzzy Sets in Algebraic Structures

The general scheme of *a-p-v* operators used in the fuzzy-set processing is presented in this section. *Lattice-like* triangular norms operate on a space of unipolar membership functions, and on the other hand, *a-p-v* operators are bipolar membership functions. To preserve compatibility with *lattice-like* operators, i.e. to have arguments and results of new operators kept in the unit interval rather than $[-1, 1]$, the space $G(X)$ and the isomorphic transformation from $F(X)$ to $G(X)$ are introduced. This transformation gives a view on merit features of new operators compatible with the classical approach. So, one can compare results of classical triangular norms and *a-p-v* operators applied to a given task.

The proposed scheme of the fuzzy set processing may be expressed in the following steps (see also Fig. 1(a)):

- transformation of the fuzzy space $F(X)$ to the fuzzy space $G(X)$ using an isomorphic mapping i,
- processing the fuzzy sets in the fuzzy space $G(X)$ by applying *a-p-v* operators,
- transforming the results back to the fuzzy space $F(X)$ using the inverse isomorphic mapping i^{-1}.

Processing in the fuzzy space $G(X)$ may be expressed in the following steps (see also Fig. 1(b)):

- transformation of the fuzzy space $G(X)$ to the respective algebraic structure using the (isomorphic) transforming function h,
- application of the usual algebraic operators,
- transforming the results back to the fuzzy space $F(X)$ using the inverse function h^{-1}.

The *a-p-v* operators introduced above are defined on the fuzzy sets whose membership values are not crisp, i.e. absolute values are less than 1. The crisp values (i.e. -1 and 1) need a special treatment because inclusion of these values breaks regularity in the linear space and ring structure of fuzzy sets. Two different attempts are proposed here to cope with fuzzy sets attaining crisp values as the grades of membership of given elements of the universe. The first attempt is based on the extension of the real line by inclusion of infinity values, i.e.

$$\underline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$$

The addition operator in $\underline{\mathbb{R}}$ will be defined accordingly for any real a:

- $a + (\pm\infty) = (\pm\infty) + a = \pm\infty$
- $(+\infty) + (+\infty) = +\infty$

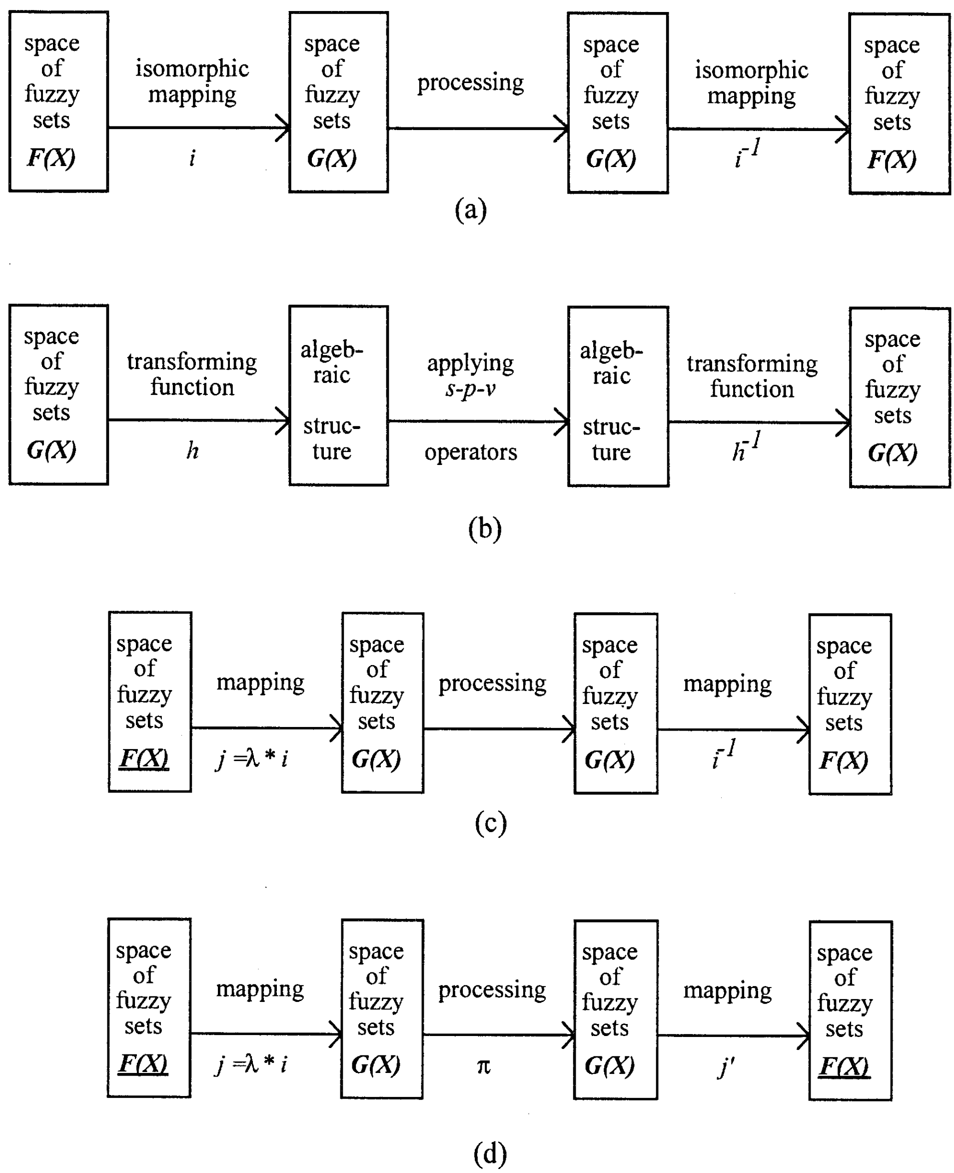

Fig. 1. Schemes of fuzzy sets processing.

Examples of the transforming function are: hyperbolic arc tangent, $x/(1-\text{abs}(x))$, $\tan(\pi x/2)$, etc. We refer the interested reader to (Homenda and Pedrycz, 1995) for discussion of the problem of selection of transforming function.

Now let a binary operator (additive operator) be such that:

$$a : G(X) \times G(X) \to G(X)$$

and for all x within the universe of discourse X:

$$a(A, B)(x) = \big(A\, a\, B\big)(x) = h^{-1}\Big(h\big(A(x)\big) + h\big(B(x)\big)\Big)$$

where A, B are fuzzy sets, $+$ is the addition operator on real numbers, and $\times$ is the Cartesian product.

Proposition 1. *The pair $\big(G(X), a\big)$ is a commutative group.*

Define another operator v (outer multiplicative operator) such that

$$v : \mathbb{R}^1 \times G(X) \to G(X)$$

and for all real λ and all x within the universe of discourse X:

$$v(\lambda, B)(x) = \big(\lambda\, v\, B\big)(x) = h^{-1}\Big(\lambda * h\big(B(x)\big)\Big)$$

where B is a fuzzy set, $*$ is the multiplication operator on real numbers and $\times$ is the Cartesian product. For both a and v, we have the following result:

Proposition 2.

(1) $\quad (\lambda * \delta)\, v\, A = \lambda\, v\, (\delta\, v\, A),$

(2) $\quad (\lambda + \delta)\, v\, A = (\lambda\, v\, A)\, a\, (\delta\, v\, A),$

or, if we assume that v is of higher precedence than a, this yields

$$(\lambda + \delta)\, v\, A = \lambda\, v\, A\, a\, \delta\, v\, A,$$

(3) $\quad \lambda\, v\, (A\, a\, B) = \lambda\, v\, A\, a\, \lambda\, v\, B,$

(4) $\quad 1\, v\, A = a.$

Finally, we conclude:

Corollary 1. *The structure $W = \big(G(X), \mathbb{R}, a, v\big)$ is a vector space.*

For proofs of the above results see (Homenda and Pedrycz, 1991).

2.2. Ring Structure of Fuzzy Sets

Let us define new operators on the space $G(X)$ of fuzzy sets based on the transforming function h introduced above. The definition of the additive operator a is adopted from the previous subsection. The definition of the (inner) multiplicative operator is as follows:

$$p : G(X) \times G(X) \to G(X)$$

and for all x within the universe of discourse X:

$$p(A,B)(x) = (A\,p\,B)(x) = h^{-1}\Big(h\big(A(x)\big) * h\big(B(x)\big)\Big)$$

where A, B are fuzzy sets, $*$ is the multiplication operator on real numbers and $\times$ is the Cartesian product.

Proposition 3. *The pair $\big(G(X), p\big)$ is a commutative group.*

Proposition 4. *The operator p is distributive with respect to the operator a:*

$$A\,p\,(B\,a\,C) = (A\,p\,B)\,a\,(A\,p\,C)$$

or, if we assume that p is of higher precedence than a, this yields

$$A\,p\,(B\,a\,C) = A\,p\,B\,a\,A\,p\,C$$

Corollary 2. *The structure $R = \big(G(X), a, p\big)$ is a ring.*

One could easily prove the following properties:

- $(A\,a\,B)(x) \geq \max\{A(x), B(x)\}$ if $A(x) \geq 0$ and $B(x) \geq 0$
- $(A\,a\,B)(x) \leq \min\{A(x), B(x)\}$ if $A(x) \leq 0$ and $B(x) \leq 0$
- $A(x) \leq (A\,a\,B)(x) \leq B(x)$ if $A(x) \leq 0 \leq B(x)$
- $(A\,p\,B)(x) \geq A(x)$ if $A(x) \geq 0$ and $B(x) \geq h^{-1}(1)$
- $(A\,p\,B)(x) \geq \max\{|A(x)|, |B(x)|\}$ if $A(x) \geq h^{-1}(1)$ and $B(x) \geq h^{-1}(1)$
- $|(A\,p\,B)(x)| \leq \min\{|A(x)|, |B(x)|\}$ if $|A(x)| \leq h^{-1}(1)$ and $|B(x)| \leq h^{-1}(1)$

It is necessary to recall that the certainty factor is interpreted here as a number belonging to the interval $(-1, 1)$ of real numbers rather than to the interval $(0, 1)$. As usual, the higher the positive value of c.f., the more certain information on inclusion of an element to the fuzzy set. And vice versa, the higher the negative value of c.f., the more certain information of excluding an element from the fuzzy set. The zero value means lack of information whether an element belongs to the given fuzzy set (Homenda and Pedrycz, 1991).

It can easily be seen that both structures are different due to different multiplicative operators. In the first case, arguments of multiplicative operator come from different spaces: the first one is a real number and may be interpreted as an importance factor, the second one is from the interval $[0, 1]$ and is simply a membership value. In the second case, both arguments come from the interval $[0, 1]$ and are of the same character—they may be relevantly interpreted as membership values.

- $(-\infty) + (-\infty) = -\infty$
- $(+\infty) + (-\infty)$ and $(-\infty) + (+\infty)$ are undefined.

For multiplication in $\mathbb{R}$ we introduce the following:

- $a * (\pm\infty) = (\pm\infty) * a = \pm\infty$ if $a > 0$
- $a * (\pm\infty) = (\pm\infty) * a = \mp\infty$ if $a < 0$
- $0 * (\pm\infty) = (\pm\infty) * 0 = 0$
- $(+\infty) * (+\infty) = (-\infty) * (-\infty) = +\infty$
- $(+\infty) * (-\infty) = (-\infty) * (+\infty) = -\infty$

Having such an extension of the real line, we can extend the definition of the transforming mapping to the closed interval $[-1, 1]$. But this involes extension of the ring and vector space to the infinity values of opposite signs. This fact has a clear interpretation: if we consider the values of the extended real line resulting from the transforming mapping, the infinity values are obtained from crisp values of the fuzzy concept. On the other hand, such crisp values of the fuzzy concept are interpreted as certain, not fuzzy, pieces of information. From this point of view, two certain but contradictory pieces of information should produce undefined results, and such a result should not be utilized in further processing as the source of this type of information cannot be viewed as reliable.

The second attempt to rectify this irregularity is based on decreasing the certainty of incoming information. It may simply be done in the step of transformation of the space of fuzzy sets $\underline{F}(X)$ into the space of fuzzy sets $G(X)$. A new mapping is obtained by multiplication of the isomorphism i by any number λ from the unit interval $(0, 1)$ (see the scheme in Fig. 1(c)). Specifically, the new mapping j is defined as follows:

$$j : \underline{F(X)} \to G(X)$$

and

$$j(x) = \lambda * (2x - 1)$$

In some cases, the above processing scheme may not be relevant to the given problem. The matter is that the output fuzzy information belongs to the space $F(X)$, i.e. the output fuzzy information does not contain crisp values, while input information is from $\underline{F(X)}$. More precisely, for each fuzzy set A from the space $F(X)$, if it is not changed in the processing step π in the space $G(X)$, the output result should be equal to A. According to problem specific features one can apply the mapping from $G(X)$ into $F(X)$ *similar* to that from $\underline{F(X)}$ into $G(X)$:

$$j'(x) = \max\Big(-1, \min\big(1, i^{-1}(x)\big)\Big) = \max\Big(-1, \min\big(1, (x+1)/2\big)\Big)$$

Note that the above mapping is not the inverse of j, but the condition

$$j'(j(A)) = A$$

is preserved for any fuzzy set A from the space $\underline{F(X)}$. The new scheme is shown in Fig. 1(d).

3. Dealing with Repetitive and Negative Information

The general meaning of repetitive and negative information is explained in Introduction. An initial discussion of unsuitability of the classical approach for these kinds of information is outlined there. In contrast with the classical approach (i.e. with the approach utilizing minimum and maximum operators or, more generally, triangular norms and conorms), application of algebraic operators to repetitive and negative information processing is presented in this section.

We refer here to Example 2, which, although rather simple, seems informative and explaining the nature of the proposed *a-p-v* operators. Of course, this example may be considered as a projection of a real-life problem on a formal space of fuzzy sets over the real line $\mathbb{R}^1$, e.g. when a market expert advises that *in order to earn a much greater profit than this year it is necessary to invest at the level of ... in advertizing the product and to invest in the research on the product at the level of ...*, he uses fuzzy opinions: *much greater* and *at the level of.* While assuming that the notion *much greater than 1* is subjectively characterized, the generalizations of given characterizations of that notion are done by processing fuzzy sets related to those characterizations. The processing of fuzzy sets is done in three formal environments: in the max-min lattice of fuzzy sets, in the triangular norm structure of fuzzy sets and in the *a-p-v* operators' structures introduced here. All these characterizations of the notion *much greater than 1* can be considered as *repetitive information.* On the other hand, the membership grades less than 0.5 can be regarded here as *negative information* according to the interpretation given in Introduction.

Example 3. Three fuzzy sets describing the notion *much greater than 5* were defined and they are given in Fig. 2(a). The noninteractivity drawback discussed in Introduction can be easily observed in Fig. 2(b), where fuzzy sets were processed by applying max-min operators and triangular norms. More specifically, distinguished characteristics shown in Fig. 2(b) were obtained by computing $\max(A(x), B(x), C(x))$, $s(A(x), B(x), C(x))$, etc., for every real x. As an example of triangular norms, the probabilistic sum and product were used.

As regards max-min operators, the non-interactivity can easily be observed: the results of max-min operators, applied to fuzzy sets A and C, are equal to the same fuzzy sets, respectively. If triangular norms are utilized, the results depend on all their arguments, so the noninteractivity problem is removed.

Nevertheless, as was mentioned before, one can expect that when membership grades are all small—less than 0.5 or even *close* to 0—the s-norm should produce a result even smaller than the arguments of that operation. This means that when

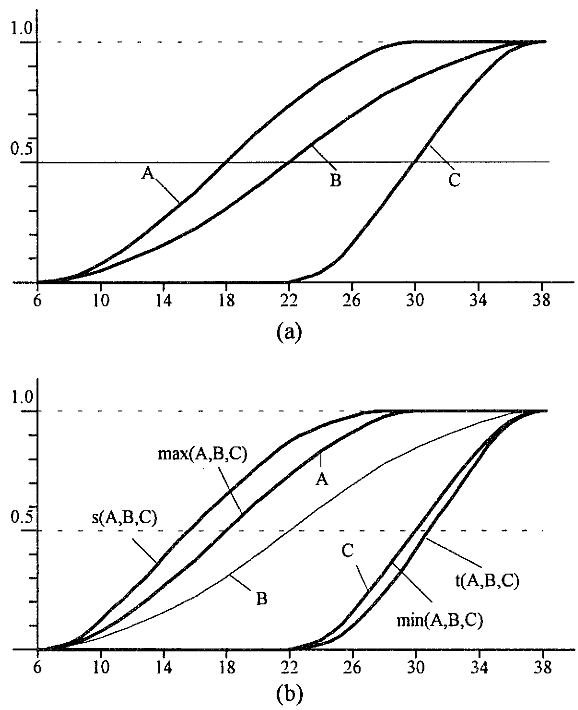

Fig. 2. Fuzzy sets processing with triangular norms.

negative pieces of information are processed, the result should not be *positive*. In other words, if all pieces of information suggest exclusion of an element from a fuzzy concept, then aggregated information should not suggest inclusion of that element into the given fuzzy concept.

As a counterexample of this expectation it is sufficient to compute the result of the probabilistic sum of two arguments both equal to 0.4. The result is then equal to 0.64. The application of the additive operator a—defined by the transforming function $h(x) = x/(1 + \text{abs}(x))$—gives the result equal to 0.33. A similar situation can be observed in Fig. 2(b).

Remark. In this paper, the probabilistic sum was used as an example of the s-norm, and the product as the t-norm. A-p-v operators were defined on the basis of the transforming function $h(x) = x/(1 + \text{abs}(x))$.

On the other hand, when the t-norm is applied to the arguments which all are greater than 0.5, the expected result should be, in any way, greater than 0.5. When applying the product of two arguments both equal to 0.6, one obtains 0.36, but the operator a gives the result equal to 0.67 (see also Fig. 2(b)). This behaviour clarifies why a-p-v operators other than lattice-like operators (the min-max operators and triangular norms) can successfully be used in modelling the interaction between items under aggregation.

Aggregation of data with s-t norms does not provide any control over the aggregation characteristics. For example, when the probabilistic sum is applied to two arguments, both equal to 0.4, the result is equal to 0.64. It is necessary to change the applied s-norm for changing this local result. In other words, it is necessary to change the s-norm applied for changing the characteristics of data aggregation.

On the other hand, a-p-v operators are provided with the feature of controlling data aggregation. This is done as a kind of local transformation of the data over aggregation. Namely, before grades of membership of different fuzzy sets are aggregated by applying the operator a, they are locally transformed by the operation p or v, and then the aggregation result may be transformed back by an *inverse-like* operation p or v.

To be more specific, the control over data aggregation can be kept in two ways:

- by controlling the parameter λ of the processing scheme given in Figs. 1(c) and 1(d) (cf. Figs. 3(a) and 3(b));
- by controlling the scalar argument of the operator v or, which is equivalent, by controlling one argument of the operator p; in the case of the operator p, additional assumptions should be made: the operator p is applied to two arguments and one of them, the non-controlled one, supports the aggregated data, while the other argument, the controlled one, influences the aggregation characteristic (cf. Figs. 3(a) and 3(b)).

Figures 3 and 4 provide an illustration for the aggregation controlling feature.

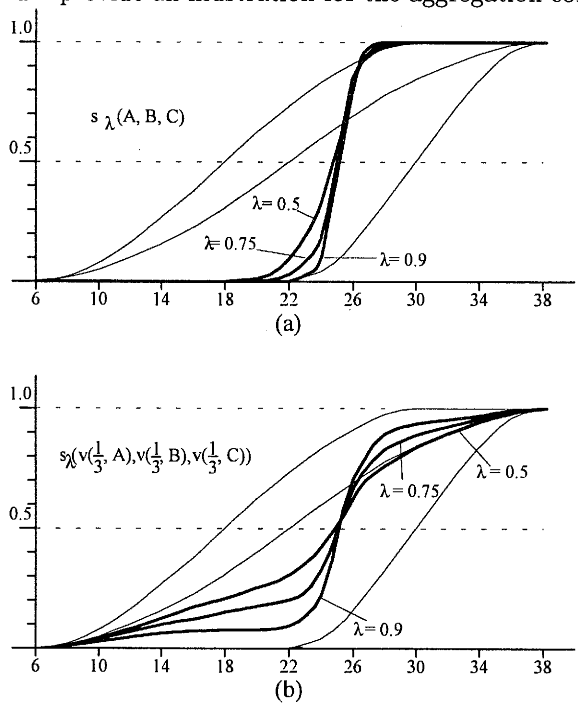

Fig. 3. Algebraic operators: crisped aggregation of fuzzy sets.

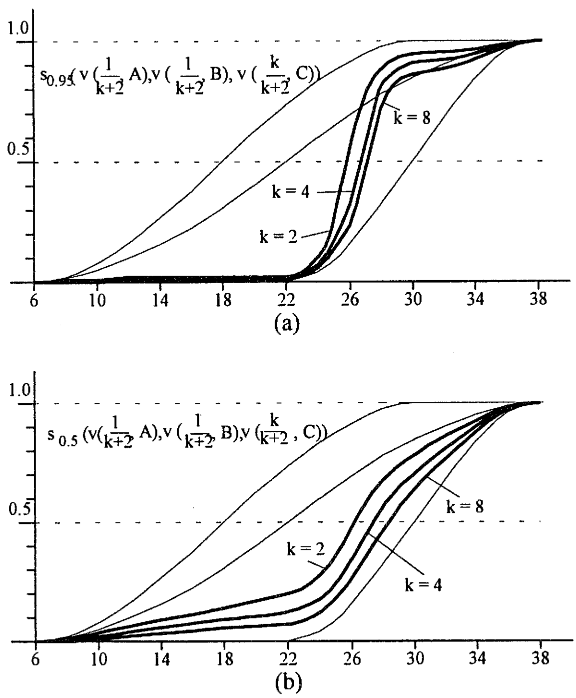

Fig. 4. Algebraic operators: weighted aggregation of fuzzy sets.

By picking up freely any of the parameters mentioned above, one can develop parametrically a broad spectrum of the aggregation characteristics without changing *a-p-v* operators. This feature may be used for defining the importance of data being aggregated as in Fig. 4, in which the parameter k reflects the importance of fuzzy information.

4. Fuzzy Reasoning with Negative Information

Since the fundamental concept of fuzzy reasoning given by Zadeh (1968; 1971), the problems of fuzzy reasoning, fuzzy relational equations and their applications have been extensively developed e.g. in (Baldwin, 1979a; 1979b; Pedrycz, 1982; 1983; 1990; 1993; Sunchez, 1979; Yager, 1980) and others. The problem of negative information in the aspect of fuzzy reasoning was primarily discussed in (Homenda, 1994; 1991). In this section, a synthesis of this problem is provided. Advantages from the application of the *a-p-v* operators in fuzzy relational equations are also pointed out.

Let us consider the fuzzy reasoning rule of the form:

If	X is A_1	then	Y is B_1
If	X is A_2	then	Y is B_2
	....		
If	X is A_k	then	Y is B_k
	X is A_0		
			Y is B_0

where X and Y are names of variables, $A_0, A_1, \ldots, A_k$ and $B_0, B_1, \ldots, B_k$ are fuzzy sets in a finite universe of discourse representing the values of these variables, i.e. $A_0, A_1, \ldots, A_k$ are elements of the space $F(\{x_0, x_1, \ldots, x_k\})$, $B_0, B_1, \ldots, B_k$ are elements of the space $F(\{y_0, y_1, \ldots, y_k\})$. This form of fuzzy reasoning rule is similar to that given by an expert in the knowledge acquisition process in a diagnostic model. Diagnostic models are of particular interest since they cope with uncertainty existing in many real-life cases concerning either a medical diagnosis or a diagnosis of technical devices. For example, fuzzy concepts $A_1, \ldots, A_k$ can be considered as coming from the space of symptoms and fuzzy concepts $B_1, \ldots, B_k$ from the space of faults (diseases). Descriptions obtained from the expert are of the form: if symptoms $x_1, \ldots, x_k$ are observed with certainty factors $a_{i1}, \ldots, a_{ik}$, then faults $y_1, \ldots, y_k$ should be considered with certainty factors $b_{i1}, \ldots, b_{ik}$, respectively. Having a given space of symptoms $x_1, \ldots, x_k$ with certainty factors $a_1, \ldots, a_k$, we face the problem of concluding on the relevant space of diseases $y_1, \ldots, y_k$ and on their certainty factors $b_1, \ldots, b_k$. This form of fuzzy reasoning may also be expressed as the fuzzy relational equation

$$X \sigma R = Y \tag{1}$$

where R is a fuzzy relation obtained from the above fuzzy reasoning rule, e.g. by the application of the methods described in works mentioned below.

By interpreting the composition operator σ in (1) as a max-min composition, e.g.

$$Y(y) = \max \big(X(x) \wedge R(x,y)\big)$$

one could easily check some basic results:

- the existence of a solution to eqn. (1) with X or R unknown (Pappis and Sugeno, 1985; Pedrycz, 1982; 1990),
- the structure of the set of solution (Di Nola, 1983),
- a solution algorithm (Pedrycz, 1990).

Unfortunately, this approach, though theoretically well-explored, has also a *non-interactivity* drawback. In this case, let us discuss the problem of noninteractivity and negative information by exploring one more example. It sheds light on the differences between max-min operators, *s-t* norms and *a-p* operators. The *s-t* norms and *a-p* operators used here are the same as in Example 3, i.e. the probabilistic sum is taken as the s-norm, the product as the t-norm, and the function $h(x) = x/(1 - \text{abs}(x))$ as the transforming function for defining the a-p operators.

Example 4. Let fuzzy relation equal

$$R = \begin{matrix} 0.6 & 0.75 & 0.8 \\ 0.7 & 0.75 & 0.6 \\ 0.75 & 0.65 & 0.8 \\ 0.7 & 0.7 & 0.55 \end{matrix}$$

while the fuzzy sets of symptoms are equal to

$$A_1 = [0.2\ 0.6\ 0.5\ 0.6]$$

$$A_2 = [0.55\ 0.6\ 0.5\ 0.6]$$

$$A_3 = [0.8\ 0.75\ 0.6\ 0.7]$$

$$A_4 = [0.95\ 0.9\ 0.95\ 0.9]$$

Fuzzy sets of faults obtained from equations $A \sigma R = B$ are shown in Table 2.

Tab. 2. Results obtained from (1).

	max-min	s-t norms	s-p operators
$A_1 \sigma R$	[0.6 0.6 0.6]	[0.82 0.82 0.78]	[0.48 0.24 0.16]
$A_2 \sigma R$	[0.6 0.6 0.6]	[0.86 0.87 0.86]	[0.63 0.67 0.60]
$A_3 \sigma R$	[0.75 0.75 0.8]	[0.94 0.95 0.95]	[0.85 0.88 0.90]
$A_4 \sigma R$	[0.75 0.75 0.8]	[0.98 0.99 0.99]	[0.97 0.98 0.98]

The max-min operators produce the same faults for both pairs of symptoms: A_1, A_2 and A_3, A_4. The application of s-t norms removes this disadvantage—fuzzy sets of faults are differentiated. In this case, membership values can be seen as much higher than practically expected, e.g. compare the vector of symptoms $A_1 = [0.2, 0.6, 0.5, 0.6]$ with the resulting fuzzy set of faults $B_1 = [0.82, 0.82, 0.78]$, and the vector of symptoms $A_2 = [0.55, 0.6, 0.5, 0.6]$ with the resulting fuzzy sets of faults $B_2 = [0.86, 0.87, 0.86]$. Let us recall that membership values *close to 0.5* are interpreted as providing little information about the inclusion of the respective element of the universe into a fuzzy concept and that a value *close to 1* means that an element of the universe is almost certainly included in the fuzzy concept, while a value *close to 0*—that the corresponding element is almost surely excluded from the fuzzy concept. For this interpretation of membership grades, the results obtained by applying s-t norms to the vector of symptoms A_1 are too high. Notice that only the membership grade of the first symptom (*close to 0*) means that this symptom should be excluded from the set of symptoms under consideration, grades of other symptoms are *close to 0.5* and our knowledge about their inclusion/exclusion is rather poor. This is the reason why we can expect here less certain knowledge about faults than if we had applied *s-t* norms.

On the other hand, the use of *a-p* operators brings expected results. In particular, very interesting results are obtained when the vector of symptoms A_1 is processed. The obtained results $[0.48, 0.24, 0.16]$ can be easily explained. The decrease in membership values of faults was caused by the increase in values in the first row of the fuzzy relation. This row is related to the first element of A_1, which, in fact, brings *negative* information. The *negative* information is *cumulated* because the respective element of the relation increases. Of course, there is some influence from the second and fourth rows of the fuzzy relation on the resulting fuzzy set of faults, but no influence comes from the second row of the fuzzy relation. This lack of influence of the second row is due to the respective (third) element of the fuzzy set of symptoms, which is equal to 0.5, which means that it brings no information.

5. Modelling an Artificial Neuron with Algebraic Operators

In this section, a fuzzy neural network is explored from the perspective of the task performed by a single neuron in one network pulse. Since the computational model of a fuzzy neuron involves classical fuzzy operations such as triangular norms, it still suffers from drawbacks outlined in the previous sections, which are related to the negative and repetitive character of input signals being processed by the neuron. To avoid these disadvantages, a model of an *algebraic neuron* is proposed on the basis of *a-p-v* operators.

The architecture of a (classical and fuzzy) neuron is shown in Fig. 5 (see also (Bezdek, 1992a; 1992b; Gupta, 1994)).

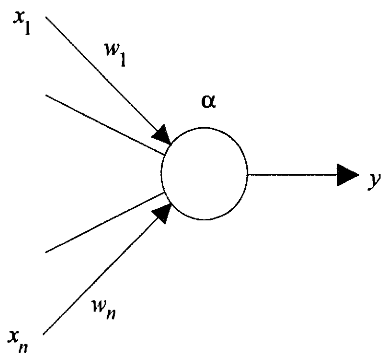

Fig. 5. Architecture of classical and fuzzy neuron.

From the mathematical point of view, processing information within a neuron involves two distinct mathematical operations, namely:

- an integration function (a synapsic operation) f first integrates the (synapsic) weights $W = (w_1, \ldots, w_n)$ with pulse input $X = (x_1, \ldots, x_n)$, and is followed by
- a transfer (activation) function (a somatic operation) F applied to the value $y = f(x)$.

The integration function f is usually the inner product:

$$f(X) = w_1 x_1 + \ldots + w_n x_n + \alpha$$

where α is the bias or offset from the origin of $\mathbb{R}^n$ to the hyperplane normal to W defined by the equation:

$$w_1 x_1 + \ldots + w_n x_n + a = 0$$

This model of an artificial neuron is called the first-order neuron because f is an affine (linear when $\alpha = 0$) function of its input, cf. (Bezdek, 1992b). Although more complicated models of neurons have been developed, including higher-order neurons, e.g. neurons with quadratic integration function, the attention is focused here on this simple, first-order neuron because this model is the basis of the artificial fuzzy neuron.

The activation function F is used to decide whether the node should fire and, if so, how much charge, and of what sign, should be broadcast to the network in response to the node inputs. F is typically the logistic (sigmoidal) function, e.g. $F(z) = F(f(X)) = 1/(1 + \exp(-z))$.

One of the models of fuzzy neurons is defined by replacing the integration function f with triangular norms:

$$u(t) = s(w_1\, t\, x_1, \ldots , w_n\, t\, x_n)$$

where t and s are the triangular norm and co-norm, respectively.

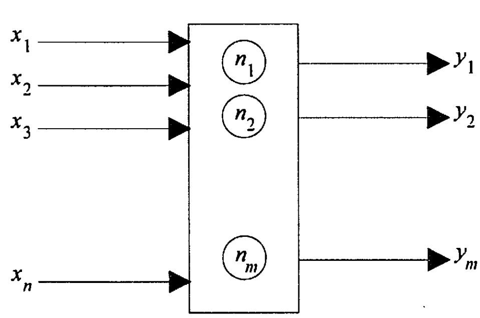

Fig. 6. Architecture of neural network layer.

With respect to one layer of the fuzzy neural network (see Fig. 6), this model of the fuzzy neuron leads to a fuzzy relational equation as a mathematical model of the layer. It may be assumed that every input of the layer is copied the number of times equal to the number of neurons in this layer. Thus, each input of the layer comes to each neuron of this layer. If this assumption is not true, the network may be extended by adding new connections going to the corresponding neurons of the layer with weights equal to 0. One can label all input connections by putting two indices for every weight: the first index is equal to the number of input connections, the second—to the number of neurons receiving the connections. In this way, all the weights of the layer may be arranged in the table:

$$W = \begin{bmatrix} w_{11} & \cdots & w_{1m} \\ & \cdots & \\ w_{n1} & \cdots & w_{nm} \end{bmatrix}$$

The output of the layer can be described by an equation similar to (1):

$$X \, \sigma \, W = Y \tag{2}$$

This result shows that issues discussed in previous sections are applicable to the model of fuzzy neural networks explored here. Specifically, the problems of noninteractivity, handling repetitive and negative information, selection of proper s-t norms, control over data aggregation, and so on, should be considered in this model of the fuzzy neural network.

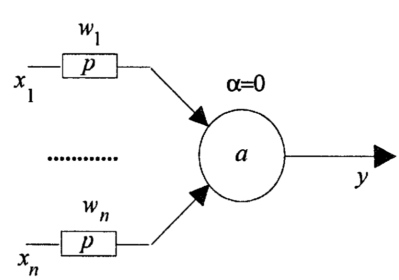

Fig. 7. Architecture of algebraic neuron.

The a-p-v operators defined and discussed in this paper seem to be a natural solution for handling negative information, defining connection weights, controlling data aggregation by integration function, describing the task of one algebraic neuron as well as of the whole layer and the whole network in one pulse. The architecture of an algebraic neuron is given in Fig. 7. This neuron aggregates inputs $x = (x_1, \ldots, x_n) \in (-1, 1)$ by first combining them individually at the synaptic level with connections (weights) $w = (w_1, \ldots, w_n) \in (-1, 1)$ and, afterwards, globally aggregating this results. The discussion of this model is restricted here to the open interval $(-1, 1)$. The values -1 and 1 may be regarded as asymptotic behaviour of respective operators. An extended discussion of this topic is going to be presented in forthcoming papers. Figure 7 illustrates the fact that the synaptic operation is in general a nonlinear operation applied to the signals and weights. The mathematical model of this neuron is summarized as follows:

$$y = A_{i=1}^{n}(w_i\, p\, x_i)$$

$$z_i = w_i\, p\, x_i = p(w_i\, p\, x_i) = f^{-1}\Big(f(w_i) * f(x_i)\Big)$$

$$y = A_{i=1}^{n}(z_i) = f^{-1}\Big(\sum_{i=1}^{n} f(x_i)\Big) = f^{-1}\Big(\sum_{i=1}^{n} (f(w_i) * f(x_i))\Big)$$

From the functional point of view, the algebraic neuron may be regarded as an intermediate model that can be situated between the classical and fuzzy neuron. Similarities between the algebraic and classical neurons can easily be shown: when classical neuron inputs and weights are normalized to $(-1, 1)$, the difference is only in the way the inputs and weights are combined, which is nonlinear in the algebraic model and the simple product in the classical model. The inverted transforming function, f^{-1}, plays here a role similar to the activation function.

On the other hand, an algebraic a-operator, when its domain is restricted to the unit square, is an Archimedean s-norm and, because of this, the activation function of the algebraic neuron (on a restricted domain of the function) is exactly equal to the activation function of the respective fuzzy s-t neuron. As to the synaptic operation, when the input and weight domains are restricted to the unit square, the product of (nonlinearly transformed) input and weight may be considered as a triangular norm applied to both. Thus, the algebraic neuron can be regarded as an extension of a fuzzy s-t neuron with respect to inputs and weights with its full compatibility for nonnegative values. This extension is supposed to give much more flexibility and adaptability in a single neuron and neural network behaviour, as well as in the training process. We refer the reader to (Homenda and Pedrycz, 1995) for discussion of fuzzy and algebraic neuron behaviour for the two-dimensional case (two inputs).

6. Conclusions

This paper deals with the processing of uncertain information in the aspect of its repetitiveness and negativity. Neither of these features is handled in the lattice structure of fuzzy sets. In the case of repetitive information, the max-min operators are *noninteractive.* Triangular norms, in contrast to max-min operators, cope with repetitive information. Nevertheless, negativity still creates problems in *lattice-like* structures (structures with max-min operators and s-t norms). An alternative approach to uncertain information processing is presented in this paper. This approach is based on the extension of fuzzy sets created by new—arithmetic rather than logic—operators, called a-p-v operators. The a-p-v operators create the linear space of fuzzy sets and the ring structure of fuzzy sets instead of the lattice structure created by max-min operators or the *lattice-like* structure created by triangular norms. Both repetitive and negative kinds of information are handled in the structures introduced in the paper.

The *a-p-v* operators are applied to two example problems: fuzzy reasoning and fuzzy neural networks. Those two topics are similar from the mathematical point of view. Both can be modelled by fuzzy relational equations and, because of this, max-min operators and triangular norms applied in those two topics create similar problems related to repetitive and negative kinds of information. *A-p-v* operators, developed in this paper, remove these problems, and also give control over data aggregation.

The paper gives a general approach to processing uncertain information handled by *a-p-v* operators. This subject is in its infancy and needs detailed studies in both theoretical and practical aspects. It seems that *a-p-v* operators give a new perspective on the uncertainty. The question as to how fertile this perspective is could be answered only if further studies on this subject were performed.

Acknowledgements

The author would like to thank Professor Witold Pedrycz for his comments on the manuscript of this paper.

References

Baldwin J.F. (1979a): *A new approach to approximate reasoning using fuzzy logic.* — Fuzzy Sets & Systems, v.2, pp.309–325.

Baldwin J.F. (1979b): *Fuzzy logic and its application to fuzzy reasoning*, In: Advances in Fuzzy Set Theory and Applications (Gupta M.M., Ragade K.R. and Yager R.R., Eds.). — Amsterdam: North-Holland, pp.83–115.

Bezdek J.C. (1992a): *Computing with uncertainty.* — IEEE Communication Magazine, pp.24–36.

Bezdek J.C. (1992b): *On the relationship between neural networks, pattern recognition and intelligence.* — Int. J. Approximate Reasoning, v.6, pp.85–107.

Dubois D.D. and Prade H. (1980): *Fuzzy Sets and Systems, Theory and Application.* — New York: Academic Press.

Gupta M.M and Rao D.H. (1994): *On the principles of fuzzy neural networks.* — Fuzzy Sets and Systems, v.61, pp.1–18.

Homenda W. (1991): *Fuzzy relational equations with cumulative composition operator as a form of fuzzy reasoning.* — Proc. Int. *Fuzzy Engineering* Symp. 91, Yokohama, pp.277–285.

Homenda W. (1994): *Fuzzy reasoning with feature cumulation as cluster analysis.* — Proc. World Congress *Computational Intelligence*, Orlando, June, pp.1922–1926.

Homenda W. and Pedrycz W. (1991): *Processing of uncertain information in linear space of fuzzy sets.* — Fuzzy Sets & Systems, v.44, pp.187–198.

Homenda W. and Pedrycz W. (1995): *Fuzzy neuron modelling based on algebraic approach to fuzzy sets.* — Proc. SPIE's Int. Symp. *Applications of Fuzzy Logic Technology II*, Orlando, Florida, SPIE v.2493, pp.71–82.

Kosko B. (1986): *Fuzzy cognitive maps.* — Int. J. Man-Machine Studies, v.24, pp.65–75.

Di Nola A. and Sessa S. (1983): *On the set of solutions of composite fuzzy relation equation.* — Fuzzy Sets & Systems, v.9, pp.275–286.

Pappis C.P. and Sugeno M. (1985): *Fuzzy relational equations and the inverse problem.* — Fuzzy Sets & Systems, v.15, pp.79–90.

Pedrycz W. (1982): *Fuzzy relational equations with triangular norms and their resolutions.* — BUSEFAL, v.11, pp.24–32.

Pedrycz W. (1983): *Numerical and applicational aspects of fuzzy relational equations.* — Fuzzy Sets & Systems, v.11, pp.1–18.

Pedrycz W. (1990): *Inverse problem in fuzzy relational equations.* — Fuzzy Sets & Systems, v.36, pp.277–291.

Pedrycz W. (1993): *s-t fuzzy relational equations.* — Fuzzy Sets & Systems, v.59, pp.189–195.

Sanchez E. (1979): *Medical diagnosis and composite fuzzy relations*, In: Advances in Fuzzy Set Theory and Applications (Gupta M.M., Ragade K.R. and Yager R.R., Eds.). — Amsterdam: North-Holland, pp.83–115.

Yager R.R. (1980): *An approach to inference in approximate reasoning.* — Int. J. Man-Machine Stud. v.13, pp.323–338.

Zadeh L.A. (1968): *Fuzzy algorithms.* — Inform. and Control, v.12, pp.94–102.

Zadeh L.A. (1971): *Similarity relations and fuzzy orderings.* — Inform. Sci. v.3, pp.177–200.

Zimmermann H.-J. and Zysno P. (1980): *Latent connectives in human decision making.* — Fuzzy Sets & Systems, v.4, pp.37–51.