Available online at www.sciencedirect.com

ScienceDirect

European Journal of Operational Research 193 (2009) 520–529

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/ejor

Computing, Artificial Intelligence and Information Management

Statistically grounded logic operators in fuzzy sets

W. Pedrycz *

Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6R 2G7
Warsaw School of Information Technology, Newelska 6, Warsaw, Poland

Received 22 September 2006; accepted 5 December 2007
Available online 29 January 2008

Abstract

In this study, we are concerned with a category of logic operators for fuzzy sets that is inherently associated with underlying statistical properties of membership grades they operate on. These constructs are referred to as *statistically* grounded logic operators, namely statistically grounded ORs (SORs, for short) and statistically grounded ANDs (SANDs, for short) operators. In essence, they arise as a solution to the optimization problem of the following form:

$$\arg\min_{m\in[0,1]} \sum_{k=1}^{N} w(x_k)|x_k - m|,$$

where $\{x_k\}$, $k = 1, 2, \ldots, N$ are the corresponding membership grades to be combined. The weight function w: $[0,1] \rightarrow [0,1]$ captures the nature of the logic *and* and *or* aggregation, respectively. We show that the weight function could be induced by t-norms and t-conorms. The weight functions could be also implied by some statistical characteristics of data (membership grades). The choice of the t-norm (t-conorm) depends upon the predefined form of the logic operators to be developed.

We demonstrate that SANDs and SORs offer an efficient operational framework for constructing fuzzy rough sets; lead to the increased sensitivity of computing possibility and necessity measures, bring a new insight into fuzzy relational equations and deliver an interpretation vehicle for fuzzy clustering (that is provided in the form of dependency analysis).

Keywords: Statistically grounded logic operators; Triangular norms; Statistical evidence; Median; Aggregation; Possibility and necessity measure; Fuzzy rough sets; Fuzzy clustering; Dependency analysis; Inter cluster relationships

1. Introductory notes

There is a wealth of various formal models of logic connectives in fuzzy sets. Alluding to the semantics of fuzzy sets and their fundamental logic operators, one can refer to such evident accomplishments in the area as t-norms and t-conorms [11–13,24], compensative operators [33], aggregative operators [2,4,7,25,27], ordered weighted operators, OWA [26], uninorms [8,28], and nullnorms [3]. Each of these categories provides additional functionality and in this way offers a highly desirable flexibility to cope with the existing diversity of problems in which fuzzy sets are used. There has been a long way we moved from the introduction of original lattice (min and max) operators on fuzzy sets. In spite of the progress being witnessed in the area, all the pursuits have been predominantly (if not exclusively) motivated by algebraic and logical underpinnings. Surprisingly, not the same amount of attention has been paid to the properties of logic operators and their developments pertinent to handling of numeric experimental data (and membership grades), cf. [10,33]. These issues are crucial given the need for fostering more advanced and effective techniques of fuzzy modeling. It is needless to say that further advancements in the development of fuzzy systems along with their further applications have posed significant modeling challenges both at the conceptual as well as the optimization end. To address them, there is a definite need for more advanced and computationally plausible logic operators.

* Present address: Warsaw School of Information Technology, Newelska 6, Warsaw, Poland.
E-mail address: pedrycz@wsiz.edu.pl

0377-2217/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2007.12.009

In particular, parameterized versions of logic operators are of interest as they bring to system modeling the highly desirable flexibility that becomes a genuine necessity when dealing with experimental data. In spite of the number of accomplishments in the realm of the fabric of the logic operators, there are still open questions that deserve careful attention. This concerns issues dealing with a non-pointwise (localized) nature of fuzzy set connectives, cf. [17] and a carefully organized mechanism of incorporation of statistical evidence into logic operators.

The problem we are interested in this study relates to the following scenario. Let us consider a finite collection of membership grades, denoted here by $x_1, x_2, \ldots, x_N \in [0,1]$ that are subject to some *and* or *or* aggregation. If the number of arguments (N) involved in either of this aggregation increases, then the result of the operations x_1 *and* $x_2 \ldots$ *and* x_N and x_1 *or* $x_2 \ldots$ *or* x_N with the *and* and *or* operators implemented with the use of Archimedean t-norms and t-conorms, respectively, quite quickly tend to zero or one, respectively. The only exception is the idempotent minimum and maximum operators. Nevertheless they come with a certain drawback such as a complete lack of interactivity; no matter how the membership grades are distributed, we exclusively rely on the maximal and minimal values in the entire set and completely ignore the remaining components. Their distribution, which could be quite crucial, is neglected in the construct of the logic operator.

We encounter a number of constructs in which there is a significant number of arguments involved in the computing with the use of fuzzy sets. Here we offer just a few of highly representative examples:

- in system modeling, we consider architectures with a significant number of inputs being membership grades associated with the corresponding fuzzy sets. For instance, in multivariable fuzzy rule-based systems, we encounter rules of the form
  – if input$_1$ is A_1 and input$_2$ is B_1 and ... and input$_n$ is W_1 then conclusion is Z.
  where A_1, B_1, etc. are fuzzy sets defined in the corresponding spaces. The activation level of such rule is typically taken as an *and* aggregation of the activation levels (membership grades) being caused by the given vector of inputs. Definitely, if the dimensionality of the problem (number of inputs increases) then the activation level of any rule (no matter how close to 1 these activation levels are) tends quite rapidly to zero.
- Consider two fuzzy sets $A = [a_1 a_2 \ldots a_N]$ and $B = [b_1 b_2 \ldots b_N]$ defined in a certain finite space $\mathbf{X}$ where card $(\mathbf{X}) = N$. The generalized possibility measure of A and B, Poss(A,B), is defined as follows:

  $$\mathrm{Poss}(A,B) = S_{i=1}^{N}(a_i t b_i) = S_{i=1}^{N} x_i \qquad (1)$$

  with $x_i = a_i t\, b_i$ where S is a certain t-conorm taken over the successive arguments $x_1, x_2, \ldots, x_N$. The generalization is sought in terms of the use of some t-conorm in the definition of the measure. In particular, one could consider the maximum operation in which case we end up with a commonly encountered definition of the possibility measure. Given a large number of elements of $\mathbf{X}$, we could easily end up with possibility measure approaching the values close to 1. Likewise, the generalized necessity measure, Nec(A,B), comes in the form

  $$\mathrm{Nec}(A,B) = \mathbf{T}_{i=1}^{N}((1-a_i) s b_i) = \mathbf{T}_{i=1}^{N} z_i \qquad (2)$$

  with $z_i = (1-a_i) s b_i$ and $\mathbf{T}$ being a certain t-norm computed over "N" arguments. In particular, one could envision here the application of the minimum operation returning the "standard" necessity measure. The aggregation completed over "N" x_i's leads to the results that converge to zero.
- The generalization of rough sets coming in the form of so-called fuzzy rough sets involves fuzzy sets described in terms of a family of sets (forming the indiscernibility relation) and leads to the description in the form of the upper and lower approximations. In essence, these approximations are computed in the form of the possibility measure (upper approximation) and necessity measure (lower approximation). The effect of convergence of these measures to 1 and 0 shown before fully applies to this case as well. Subsequently we may easily end up with very "loose" description of the fuzzy rough set.
- In fuzzy relational calculus we typically encounter fuzzy relational equations [5,18,19] of the form $A \circ R = B$ where A and B are fuzzy sets defined in some finite spaces $\mathbf{X}$ and $\mathbf{Y}$ with R being a certain fuzzy relation defined in Cartesian product of $\mathbf{X}$ and $\mathbf{Y}$. The aggregation operator combining A and R is the one of s–t composition and max–min, in particular. The dual type of the fuzzy relational equation, $A \bullet R = B$ uses the t–s composition operator (with the min–max composition being its very special case). Again when the number of elements of $\mathbf{X}$ increases, then the results of these two compositions tend to 1 and 0, respectively.

Being alerted by this phenomenon (and its practical implications) in which we are faced with a finite (yet quite large) population of membership grades, we are at position that the logic operators have to be revisited so that they exhibit clear logic facet however in their design we take advantage of the underlying experimental evidence. In this sense, we envision that such revisited constructs could benefit when being positioned at the junction of logic and the use of the available statistical evidence (results). Our ultimate objective is to introduce logic operators whose construction seamlessly embrace the logic fabric and augment it by the existing experimental evidence. Given this, we will be referring to them as statistically grounded *OR* (SOR) and statistically grounded *AND* (SAND) logic operators. In the context of our investigations, it is worth referring to the study by Greco et al. [9] in which they offered a view at aggregation mechanisms realized in terms of gradual rules. In this sense, the approach proposed by them is complementary to the ideas developed here.