International Journal Of General System

ISSN: 0308-1079 (Print) 1563-5104 (Online) Journal homepage: https://www.tandfonline.com/loi/ggen20

ROUGH FUZZY SETS AND FUZZY ROUGH SETS*

DIDIER DUBOIS & HENRI PRADE

To cite this article: DIDIER DUBOIS & HENRI PRADE (1990) ROUGH FUZZY SETS AND FUZZY ROUGH SETS*, International Journal Of General System, 17:2-3, 191-209, DOI: 10.1080/03081079008935107

To link to this article: https://doi.org/10.1080/03081079008935107

Published online: 06 Apr 2007.

Int. J. General Systems, Vol. 17, pp. 191–209
Reprints available directly from the publisher
Photocopying permitted by license only

ROUGH FUZZY SETS AND FUZZY ROUGH SETS*

DIDIER DUBOIS and HENRI PRADE

Laboratoire Langages et Systèmes Informatiques, Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse Cedex, France

(*Received 28 November 1988; in final form 18 July 1989*)

The notion of a rough set introduced by Pawlak has often been compared to that of a fuzzy set, sometimes with a view to prove that one is more general, or, more useful than the other. In this paper we argue that both notions aim to different purposes. Seen this way, it is more natural to try to combine the two models of uncertainty (vagueness and coarseness) rather than to have them compete on the same problems. First, one may think of deriving the upper and lower approximations of a fuzzy set, when a reference scale is coarsened by means of an equivalence relation. We then come close to Caianiello's C-calculus. Shafer's concept of coarsened belief functions also belongs to the same line of thought. Another idea is to turn the equivalence relation into a fuzzy similarity relation, for the modeling of coarseness, as already proposed by Fariñas del Cerro and Prade. Instead of using a similarity relation, we can start with fuzzy granules which make a fuzzy partition of the reference scale. The main contribution of the paper is to clarify the difference between fuzzy sets and rough sets, and unify several independent works which deal with similar ideas in different settings or notations.

INDEX TERMS: Fuzzy sets, rough sets, C-calculus, random sets, belief functions, similarity relations.

1. INTRODUCTION

The contemporary concern about knowledge representation and information systems has put forward useful extensions of elementary set theory such as fuzzy sets (Zadeh[1]) and rough sets (Pawlak[2]), among others. In this paper we pursue an investigation around these two notions in order to lay bare their respective specificity, instead of turning them into rival theories (e.g. Pawlak[3]). Basically, rough sets embody the idea of indiscernibility between objects in a set, while fuzzy sets model the ill-definition of the boundary of a sub-class of this set. Rough sets are a calculus of partitions, while fuzzy sets are a continuous generalization of set-characteristic functions. Marrying both notions lead to consider rough approximations of fuzzy sets, but also approximation of sets by means of similarity relations or fuzzy partitions. These hybrid notions come up in a natural way when a linguistic category, denoting a set of objects, must be approximated in terms of already existing labels, or when the indiscernibility relation between objects no longer obeys the ideal laws of equivalence and is a matter of degree. Moreover, this attempt to mix up vagueness and approximation leads us to bring together

*This paper is based on a talk at the International Conference on Fuzzy Sets in Informatics, Moscow, September 20–23, 1988.

past works developed independently of rough sets and often before them, but based on the same ideas.

The first section contains basic definitions of rough sets and fuzzy sets and points out the contrast between the intended purposes of the two notions. Section two defines upper and lower approximations of fuzzy sets and belief functions, and bridges the gap with Caianiello's C-calculus.[4] It is indicated that under a different terminology, Shafer[5] has used the same model as Pawlak in the theory of evidence, i.e. a calculus of partitions. Section three generalizes rough sets by weakening the concept of equivalence into similarity (Zadeh[6]), as suggested in a previous paper.[7] The rough set idea then comes close to well-known concerns in mathematical taxonomy and approximation theory. This section equivalently builds fuzzy rough sets from fuzzy partitions and shows that this notion can address the problem of linguistic approximation (Zadeh,[8] Bonissone,[9] for instance). Relevant results relating fuzzy similarity relations and partitions are recalled. Basic properties of fuzzy rough sets are investigated. In the conclusion, some research directions are surveyed.

2. FUZZY SETS AND ROUGH SETS: TWO DIFFERENT TOPICS

Let X be a set, and R be an equivalence relation on X (i.e. reflexive, symmetric and transitive). Let X/R denote the quotient set of equivalence classes, which form a partition in X. X/R is a coarsened version of X; elementary parts of X/R are coarser than the ones of X, and denoted $X_1, X_2, \ldots, X_i, \ldots$. The cardinality of X/R is generally smaller than that of X. An equivalence relation is the simplest model one can think of to represent the fact that, in X, it is not possible to distinguish some elements from others. xRy then means: x is too close (or too similar) to y so that both elements are indiscernible.

Examples 1) *Measurement scale* $X = [0, 2.5]$ is a human size scale between 0 and 2.5 meters, that allows for infinite precision. In practice, only millimeters can be measured, i.e. X/R is a set of adjacent intervals, whose representatives are of the form $n/1000$ with $0 \leqq n \leqq 2500$, n integer. $xR(n/1000)$ means that x can be rounded by $n/1000$, xRy means that x and y are rounded by the same number of millimeters. In usual communication between individuals on this matter, the implicit representation of this scale is even coarser: only centimeters (or inches) make sense.

2) *Information system* X is a set of item identifiers (objects), $\mathscr{A}$ is a set of attributes a, V_a the set $\{a(x) | x \in X\}$ of attribute values for attribute a. The equivalence relation R is defined by xRy if and only if $\forall a \in \mathscr{A}$, $a(x) = a(y)$. $[x]_R$ denotes the class of objects which have the same description as x in terms of attributes in $\mathscr{A}$. This example is given by Pawlak.[10]

3) *Image processing* X is a rectangle screen, i.e. a Cartesian product $[0, a] \times [0, b]$, X/R is a discretization grid into pixels, $[(x, y)]_R$ being the pixel that contains a point (x, y) in X. This is the 2-dimensional version of example 1.

Let S be a subset of X. The main question addressed by rough sets (Pawlak[2]) is: how to represent S by means of X/R? Denote $[x]_R$ the equivalence class of $x \in X$. A *rough set* is a pair of subsets $R^*(S)$ and $R_*(S)$ of X/R that approach as close as possible S from outside and inside respectively:

$$R^*(S)=\{[x]_R \mid [x]_R \cap S \neq \varnothing, x \in X\} \tag{1}$$

$$R_*(S)=\{[x]_R \mid [x]_R \subseteq S, x \in X\} \tag{2}$$

$R^*(S)$ (resp.: $R_*(S)$) is called the upper (resp.: lower) approximation of S by R. $R^*(S)$ contains $R_*(S)$. When $R^*(S) \neq R_*(S)$, it means that due to the indiscernibility of elements in X, S cannot be perfectly described. More precisely, the set difference $R^*(S)-R_*(S)$ is a rough description of the boundary of S by means of "granules" of X/R.

These notions are actually older than Pawlak's paper. They were already introduced by Shafer[5] in his book (chapter 6), where coarsenings and refinements of a frame of discernment are introduced. A frame of discernment is a set of alternatives perceived as distinct answers to a question. Coarsening a frame of discernment X comes down to clustering elements and build a partition. Refinement is the converse operation, i.e. distinguishing sub-alternatives corresponding to single elements. In other words X/R is a coarsening of X, and X is a refinement of X/R. Following Shafer, let us denote ω the mapping that, for any subset of X/R, computes its refinement in X. Namely if $X_i \in X/R$,

$$\omega(X_i)=\{x \mid X_i \text{ is the name of the equivalence class } [x]_R\} \tag{3}$$

and for $A \subset X/R$,

$$\omega(A)=\bigcup_{X_i \in A} \omega(X_i). \tag{4}$$

It is important to distinguish between X_i, an element of X/R, and $\omega(X_i)$, a subset of X. $R^*(S)$ and $R_*(S)$ are respectively called outer and inner reductions by Shafer.[5] Viewing X_i as the name of an equivalence class, $\omega(X_i)$ can be viewed as the extension of X_i, and will be termed so in the following, by analogy with logic.

A fuzzy set[1] F of X is defined by a mapping $\mu_F\colon X \to L$ where L is an ordered set of membership values (often a complete lattice, at least) and $\mu_F(x)$ is the degree of membership of x in F. $L=[0,1]$ generally; $L=\{0,1\}$ in the case of usual sets. Allowing for partial membership intends to account for the ill-definition of the extension of the predicate named F.

Examples (continued) 1) *Measurement scale*: Rounding off sizes to millimeters does not always allow to distinguish between people whose size is close to within one millimeter, i.e. if $S=\{1.71815,\ 1.71816\}$, $R^*(S)=[1.718]_R$ but $R_*(S)=\varnothing$. Contrastedly the set of *tall* sizes is fuzzy because (even in millimeters) some sizes are compatible with *tall* to a degree that may be different from total compatibility and total incompatibility. Vagueness lies in the subset F denoted by "tall" while indiscernibility is a property of the referential itself, not of its subsets. The model of "tall" changes by modification of the set L of membership values.

2) *Information systems*: Fuzziness in information systems often lies in the formulation of queries that describe subsets of relevant objects by a flexible specification of admissible attribute values. Then the set of retrieved objects is fuzzy (e.g. Tahani[11]). Moreover the presence of ill-described objects in a data base can be expressed in terms of possibility distributions on the set V_a of attribute values; these possibility distributions are modelled by fuzzy sets, following Zadeh.[12] For instance, all we know about x is that its size $a(x)$ is "about 2

meters", with the underlying assumption that $a(x)$ is in $[1.8, 2.2]$ and the most plausible value is 2. If a user is interested in "tall objects", the response of the information system may consist of 2 fuzzy sets of relevant objects: the set of certainly relevant objects, and the (larger) set of possibly relevant objects (Prade and Testemale[13]). These two sets are nested; the corresponding pair is called a *twofold* fuzzy set.[14] Its existence is due to the presence of vague incomplete information in the data base, but not to the indiscernibility of objects.

3) *Image processing*: given a subset S of the screen defined by the contour of an object, various approximations $R^*(S)$ and $R_*(S)$ of this subset can be obtained by modifying the graininess of the discretized picture. However generally, objects on a black to white screen appear rather like fuzzy sets of the screen, due to grey levels. The object can be made more or less fuzzy by acting on the number of grey levels, i.e. contrast modification. Number of allowed levels of grey n_1 and number of pixels n_2 in the screen are unrelated parameters.

It should be clear from the above examples that rough sets and fuzzy sets are not meant to play the same role in knowledge representation problems. As a consequence they are not rival theories but capture two distinct aspects of imperfection in knowledge: indiscernibility and vagueness, that may be simultaneously present in a given application. It is then natural to combine these notions, rather than compare them from a formal point of view.

Of course, it may be tempting to identify the boundary of a rough set as containing borderline elements and then decree that the upper and lower approximations of a set S can be viewed respectively as the support (i.e., elements with positive membership) and the core (i.e., elements with complete membership) of a fuzzy set $F(S)$ defined on X/R. Then, as suggested by Pawlak,[3] elements in the boundary can have membership value 0.5. This idea is further extended by Wong and Ziarko[15] (see also Pawlak *et al.*[16]) who suggest to use the conditional probability $P(S|\omega(X_i))$ to evaluate the degree of membership $\mu_{F(S)}(X_i)$. However these views can be but partially in agreement with fuzzy set theory since they impose severe restrictions on the choice of fuzzy set-theoretic connectives (see Wygralak[17] for the three-valued logic approach, and Wong and Ziarko[15] for the probabilistic view).

2. UPPER AND LOWER APPROXIMATIONS OF GENERALIZED SETS

2.1. Approximations of Fuzzy Sets

Let X be a set, R be an equivalence relation on X and F be a fuzzy set in X. The upper and lower approximations $R^*(F)$ and $R_*(F)$ of a fuzzy set F by R are fuzzy sets of X/R with membership functions defined by

$$\mu_{R^*(F)}(X_i) = \sup\{\mu_F(x) | \omega(X_i) = [x]_R\} \tag{5}$$

$$\mu_{R_*(F)}(X_i) = \inf\{\mu_F(x) | \omega(X_i) = [x]_R\} \tag{6}$$

where $\mu_{R^*(F)}(X_i)$ (resp.: $\mu_{R_*(F)}(X_i)$) is the degree of membership of X_i in $R^*(F)$ (resp.: $R_*(F)$). $(R^*(F), R_*(F))$ is called a rough fuzzy set. These expressions derive from possibility theory (Zadeh;[12] Dubois and Prade[18]); (5) (resp.: (6)) is the degree of possibility (resp.: necessity) of the fuzzy event F, based on the (crisp) possibility distribution defined from the characteristic function of $\omega(X_i)$. To see it, note that the fuzzy extensions $\omega(R^*(F))$ and $\omega(R_*(F))$ can be defined via the extension principle,[1] as:

$$\mu_{\omega(R^*(F))}(x)=\mu_{R^*(F)}(X_i)=\Pi_i(F), \quad \forall x\in\omega(X_i) \tag{7}$$

$$\mu_{\omega(R_*(F))}(x)=\mu_{R_*(F)}(X_i)=N_i(F), \quad \forall x\in\omega(X_i) \tag{8}$$

where Π_i (*resp.* N_i) is the possibility (resp. necessity) measure whose distribution is $\mu_{\omega(X_i)}$ hereafter denoted π_i, i.e. $\Pi_i(F)=\sup_x \min(\pi_i(x), \mu_F(x))$; $N_i(F)=\inf_x \max(1-\pi_i(x), \mu_F(x))$.[18] Note that π_i is a crisp possibility distribution, i.e. $\pi_i(x)\in\{0,1\}$, and $x\in\omega(X_i)$ is equivalent to $\omega(X_i)=[x]_R$ and to $\pi_i(x)=1$; this is why (5) and (6) are the same as (7) and (8).

2.2. *The Link with C-Calculus*

If we apply (7) and (8) to a fuzzy set of the real line as "tall" in Example 1, we obtain for $\mu_{\omega(R^*(F))}$ and $\mu_{\omega(R_*(F))}$ piecewise constant functions that are used in integration theory to bracket the integral of μ_F by means of Darboux sums. Moreover, (7) and (8) are basic equations of C-calculus (Caianiello[4,19]). A composite set or C-set is a triple (χ, m, M) where $\chi=\{X_1,\ldots,X_n\}$ correspond to a partition of X, and m, M are mappings $X\to[0,1]$ such that

$$\forall x\in X, m(x)=\sum m_i\mu_{\omega(X_i)}(x)=m_i \quad \text{if } x\in\omega(X_i) \tag{9}$$

$$\forall x\in X, M(x)=\sum M_i\mu_{\omega(X_i)}(x)=M_i \quad \text{if } x\in\omega(X_i) \tag{10}$$

where $0\leq m_i\leq M_i\leq 1$, $\forall i=1,n$. Every function $f: X\to[0,1]$ defines a composite set, letting $m_i=\inf\{f(x)\mid x\in\omega(X_i)\}$ and $M_i=\sup\{f(x)\mid x\in\omega(X_i)\}$. Clearly these equations are (5) and (6) where $X/R=\chi$, and $f=\mu_F$ is the membership function of a fuzzy set. The links between fuzzy sets and C-calculus were pointed out in the first paper on C-calculus[4] and more recently by Caianiello and Ventre.[20] But it is clear that a C-set is nothing but a rough fuzzy set, i.e., a more general and, by the way, earlier notion than a rough set.

A basic operation in C-calculus is C-product (Caianiello and Ventre[21]). Given two C-sets (χ, m, M) and (χ', m', M'), a C-set is obtained, say (χ'', m'', M'') such that $\chi''=\{\omega(X_i)\cap\omega(X'_j)\mid X_i\in\chi, X'_j\in\chi'\}, m''=\min(m,m')$, $M''=\max(M,M')$. However the two latter relations, expressing a fuzzy set intersection and union respectively, do not give exact result for the rough fuzzy set, defined using χ''; indeed, if m and M are defined from f as well as m' and M', only the following inequalities hold

$$\begin{aligned}\min(m_i, m'_j)&\leq\inf\{f(x)\mid x\in\omega(X_i)\cap\omega(X'_j)\}\\ &\leq\sup\{f(x)\mid x\in\omega(X_i)\cap\omega(X'_j)\}\leq\max(M_i, M'_j)\end{aligned} \tag{11}$$

It is interesting to notice that the main applications of C-calculus are in image processing, i.e. in the setting of Example 3. However C-product of rough sets make sense in Example 2. If R_a and R_b denote the equivalence relations defined on X by attributes a and b respectively, the refined equivalence relation R_{ab} defined on X by both attributes correspond to the partition obtained by intersecting the equivalence classes for attribute a and attribute b, i.e. a C-product.. In terms of rough sets (11) writes (Pawlak[10]):

$$R_{a*}(S) \cap R_{b*}(S) \subseteq R_{ab*}(S) \subseteq R^*_{ab}(S) \subseteq R^*_a(S) \cup R^*_b(S) \tag{12}$$

2.3 Approximations of Random Sets

In Shafer's book,[5] chap. 6, preliminary results are given about coarsening and refinement of belief functions. The author assumes that a belief function is defined on X/R and computes its refinement. This work is pursued in Shafer *et al.*[22] A belief function on X is defined by a finite set $\mathscr{F}$ of non-empty subsets to which positive masses $p(S), S \in \mathscr{F}$ are allocated, so that $\sum_S p(S) = 1$. When $\mathscr{F}$ contains only a family of nested sets, the belief function is called consonant and is a necessity measure in the sense of possibility theory (Dubois and Prade[18]) i.e. the random set $(\mathscr{F}, p)$ is equivalent to a fuzzy set F with

$$\forall x \in X, \mu_F(x) = \sum_{x \in S} p(S) \tag{13}$$

μ_F is called the contour function of $(\mathscr{F}, p)$ by Shafer,[5] even when $(\mathscr{F}, p)$ is not consonant. In the consonant case, the equivalence between $(\mathscr{F}, p)$ and F via (13) comes from the fact that $\mathscr{F}$ is the set of level-cuts of F, i.e. $\mathscr{F} = \{F_\alpha | \alpha \in (0, 1]\}$ with $F_\alpha = \{x | \mu_F(x) \geqq \alpha\}$; the 1-cut, called the core of F, is not empty because $\mathscr{F}$ does not contain the empty set. Coarsening X using an equivalence relation R leads to introduce upper and lower approximations of $(\mathscr{F}, p)$ say $(R^*(\mathscr{F}), p^*)$, $(R_*(\mathscr{F}), p_*)$ on X/R, with

$$R^*(\mathscr{F}) = \{R^*(S) | S \in \mathscr{F}\}; p^*(A) = \sum \{p(S) | A = R^*(S)\} \tag{14}$$

$$R_*(\mathscr{F}) = \{R_*(S) | S \in \mathscr{F}\}; p_*(A) = \sum \{p(S) | A = R_*(S)\} \tag{15}$$

(14–15) generalize (1–2) into what can be called "rough random sets".

The consistency of the rough random sets and the rough fuzzy sets (i.e. the equivlence between (14–15) and (5–6)) is easily achieved if we invert (13). Starting from a fuzzy set F on a finite set X, letting $\alpha_1 = 1 \geqq \alpha_2 \geqq \cdots \geqq \alpha_k$ be the set of positive membership grades, the random set equivalent to F is defined by[18]

$$\mathscr{F}_F = \{F_{\alpha_i}, i = 1, k\}$$

$$p_F(F_{\alpha_i}) = \alpha_i - \alpha_{i+1}$$

with $\alpha_{k+1} = 0$, by convention. We moreover need the following:

LEMMA 1 *$\forall \alpha \in (0, 1]$, $R^*(F)_\alpha = R^*(F_\alpha)$ and $R_*(F)_\alpha = R_*(F_\alpha)$.*

Proof

$$X_i \in R^*(F)_\alpha \Leftrightarrow \Pi_i(F) = \max \{\gamma | \omega(X_i) \cap F_\gamma \neq \varnothing\} \geqq \alpha$$

$$\Leftrightarrow \omega(X_i) \cap F_\alpha \neq \varnothing \Leftrightarrow X_i \in R^*(F_\alpha).$$

This proves the first half of the lemma. Note further that denoting $F_{\bar{\alpha}}$ the strong α-cut of F,[18], i.e. $F_{\bar{\alpha}} = \{x \mid \mu_F(x) > \alpha\}$, we also have $R^*(F)_{\bar{\alpha}} = R^*(F_{\bar{\alpha}})$. The last result needed is $\overline{R^*(F)} = R_*(\overline{F})$ where $\mu_{\overline{F}} = 1 - \mu_F$ is the membership of the complement of F. This is obvious from (5) and (6). Now the remainder of the lemma is obtained by the following derivation

$$R_*(F)_\alpha = \overline{(\overline{R_*(F)})_{\overline{1-\alpha}}} = \overline{(R^*(\overline{F}))_{\overline{1-\alpha}}} = \overline{R^*((\overline{F})_{\overline{1-\alpha}})} = \overline{R^*(\overline{F_\alpha})} = R_*(F_\alpha)$$

which takes advantage of a well-known result:[18] $(\overline{F})_\alpha = \overline{F_{\overline{1-\alpha}}}$. Q.E.D.

Now the following result becomes easy:

PROPOSITION 1 *$R^*(F)$ is equivalent to $(R^*(\mathscr{F}_F), p_F^*)$; $R_*(F)$ is equivalent to $(R_*(\mathscr{F}_F), p_{F_*})$.*

For instance

$$R^*(F) \Leftrightarrow (\mathscr{F}_{R^*(F)}, p_{R^*(F)}) \Leftrightarrow (R^*(\mathscr{F}_F), p_F^*) \text{ using the lemma}$$

Indeed $\mathscr{F}_{R^*(F)} = \{R^*(F)_{\alpha_i} \mid i = 1, k\} = \{R^*(F_{\alpha_i}) \mid i = 1, k\}$. As for the masses, $p_F^*(R^*(F_{\alpha_i})) = p_{R^*(F)}((R^*(F)_{\alpha_i}) = p_F(F_{\alpha_i}) = \alpha_i - \alpha_{i+1}$. Now if $R^*(F_{\alpha_i}) = R^*(F_{\alpha_{i+1}})$ for some i, then $\mathscr{F}_{R^*(F)}$ would contain possibly non-distinct sets. Allowing for non-distinct focal sets is a matter of convention and does not affect (13) the equation that produces the membership function.

2.4. Approximations of Belief Functions

Another way of proceeding is to project the belief and plausibility functions from X to X/R directly. Let Bel and Pl be two set-functions defined on 2^X, built from $(\mathscr{F}, p)$, as follows:

$$\mathrm{Bel}(S) = \sum_{T: T \subseteq S} p(T) \text{ (belief function)}$$

$$\mathrm{Pl}(S) = \sum_{T: T \cap S \neq \varnothing} p(T) \text{ (plausibility function).}$$

Then define Bel_R and Pl_R on $2^{X/R}$ as $\mathrm{Bel}_R(A) = \mathrm{Bel}(\omega(A))$; $\mathrm{Pl}_R(A) = \mathrm{Pl}(\omega(A))$. Let Bel^* and Bel_* on $2^{X/R}$ derive from $(R^*(\mathscr{F}), p^*)$ and $(R_*(\mathscr{F}), p_*)$ respectively. It is easy to see that $\mathrm{Bel}_R = \mathrm{Bel}^*$. Indeed,

$$\mathrm{Bel}^*(A) = \sum_{B: B \subseteq A} p^*(B) = \sum_{S: R^*(S) \subseteq A} p(S) = \sum_{S: S \subseteq \omega(A)} p(S) = \mathrm{Bel}(\omega(A))$$

since $R^*(S) \subseteq A$ is equivalent to $S \subseteq \omega(R_*(\omega(A))) = \omega(A)$. Similarly, $\mathrm{Pl}^* = \mathrm{Pl}_R$.

The effect of coarsening X by means of R can be analyzed on X itself. Namely, noticing that the belief function is defined by means of the set inclusion, the presence of R leads to four possible definitions of inclusion of a subset T in a subset S, when only their upper and lower approximations are discerned: $R_*(T) \subseteq R_*(S)$; $R^*(T) \subseteq R_*(S)$, $R_*(T) \subseteq R^*(S)$, $R^*(T) \subseteq R^*(S)$, respectively denoted ${}_*\subseteq_*$, ${}^*\subseteq_*$, ${}_*\subseteq^*$, ${}^*\subseteq^*$. Their strength compares as follows

$$T^* \subseteq_* S \text{ implies } T_* \subseteq_* S \text{ which implies } T_* \subseteq^* S \tag{16}$$

$$T^* \subseteq_* S \text{ implies } T^* \subseteq^* S \text{ which implies } T_* \subseteq^* S \tag{17}$$

These concepts, that may be called "rough implications", can be compared with the usual implication. Namely ${}^*\subseteq_*$ is stronger than $\subseteq$, and $\subseteq$ is stronger than the three others. ${}_*\subseteq_*$ and ${}^*\subseteq^*$ do not compare. Based on these implications four definitions of a rough belief function are obtained on X

$$\mathrm{Bel}^*{}_*(S) = \sum_{T^* \subseteq_* S} p(T)$$

$$\mathrm{Bel}_{**}(S) = \sum_{T_* \subseteq_* S} p(T)$$

$$\mathrm{Bel}^{**}(S) = \sum_{T^* \subseteq^* S} p(T)$$

$$\mathrm{Bel}_*{}^*(S) = \sum_{T_* \subseteq^* S} p(T)$$

Notice the following equivalences. $T^* \subseteq_* S \Leftrightarrow T \subseteq \omega(R_*(S))$ and $T^* \subseteq^* S \Leftrightarrow T \subseteq \omega(R^*(S))$. They can be justified by the identities $R^*(\omega(R_*(S))) = R_*(S)$, $R^*(\omega(R^*(S)) = R^*(S)$ and the inclusion $T \subseteq \omega(R^*(T))$. As a consequence

$$\begin{aligned} \mathrm{Bel}^*{}_*(S) &= \mathrm{Bel}^*(R_*(S)) \quad ; \quad & \mathrm{Bel}^{**}(S) &= \mathrm{Bel}^*(R^*(S)) \\ &= \mathrm{Bel}(\omega(R_*(S))) & &= \mathrm{Bel}(\omega(R^*(S))) \end{aligned} \tag{18}$$

i.e. $\mathrm{Bel}^*{}_*$ and Bel^{**} derive from the belief function $\mathrm{Bel}^* = \mathrm{Bel}_R$ with underlying random set $(R^*(\mathscr{F}), p^*)$.

Dubois and Prade[23] study the concept of upper and lower belief function induced by a multiple-valued mapping. If Ω and Ω' are two sets, Γ a mapping from Ω to $2^{\Omega'} - \varnothing$, Bel a belief function on Ω, and A' a subset of Ω', two belief functions Ubel and Lbel can be defined on Ω' as follows

$$\mathrm{Ubel}(A') = \mathrm{Bel}(\{\omega \in \Omega \mid \Gamma(\omega) \cap A' \neq \varnothing\}) \text{ (upper belief function)}$$

$$\mathrm{Lbel}(A') = \mathrm{Bel}(\{\omega \in \Omega \mid \Gamma(\omega) \subseteq A'\}) \text{ (lower belief function)}$$

It is nothing than iterating Dempster's[24] construction of belief functions from probability spaces. It can be proved that Lbel $\leqq$ Ubel and that Lbel is still a belief function while Ubel is not, generally.

In this paper, it is easy to figure out that

- $\mathrm{Bel}^*{}_*$ is the lower belief function induced by Bel^* from X/R to X by means of the multiple-valued mapping ω that associates to each X_i the equivalence class $[x]_R$ such that $\omega(X_i)=[x]_R$.

- Bel** is the upper belief function induced by Bel* in the same way.

This is obvious noticing that $R^*(S)=\{X_i|\omega(X_i)\cap S\neq\varnothing\}$ and $R_*(S)=\{X_i|\omega(X_i)\subseteq S\}$. Moreover the set-function $\text{Bel}^*{}_*$ is a belief function on 2^X, while Bel** is not in spite of the misleading appearance of (18). The random set equivalent to $\text{Bel}^*{}_*$ has focal elements $\{\omega(R^*(S))|S\in\mathscr{F}\}$, and the mass allocated to $\omega(R^*(S))$ is $p^*(R^*(S))$, $\text{Bel}^*{}_*$ derives from the extension of $(R^*(\mathscr{F}),p^*)$. To see it just remember the identity between Bel_R and Bel* on X/R, that is based on the equivalence between $T\subseteq\omega(R_*(S))$ and $\omega(R^*(T))\subseteq S$, so that

$$\begin{aligned}\text{Bel}^*{}_*(S) &= \sum_{T\subseteq\omega(R_*(S))} p(T) \qquad \text{from (18)}\\ &= \sum_{\omega(R^*(T))\subseteq S} p(T)\\ &= \sum_{\omega(A)\subseteq S} p^*(A)\end{aligned}$$

These properties of $\text{Bel}^*{}_*$ have been studied by Shafer *et al.*,[22] where $\text{Bel}^*{}_*$ is called the "coarsening" of Bel, moreover Shafer *et al.*[22] notice that

$$\text{Bel}^*{}_*(S)=\max\{\text{Bel}(T)|T\subseteq S, T=\omega(A) \text{ for some } A\subseteq X/R\} \tag{19}$$

a relation that is obvious from (18). Similarly

$$\text{Bel}^{**}(S)=\min\{\text{Bel}(T)|S\subseteq T, T=\omega(A) \text{ for some } A\subseteq X/R\} \tag{20}$$

since $\omega(R^*(S))$ is the smallest subset containing S and being the extension of a subset of X/R. But Bel** is generally not a belief function on 2^X.

Let us now consider the set-functions induced by the lower approximation of $(\mathscr{F},p)$, i.e. the upper belief function $\text{Bel}_*(R^*(S))$ and the lower belief function $\text{Bel}_*(R_*(S))$. Clearly

$$\begin{aligned}\text{Bel}_*(R^*(S)) &= \sum_{A\subseteq R^*(S)} p_*(A)\\ &= \sum_{R_*(T)\subseteq R^*(S)} p(T)\\ &= \text{Bel}_*{}^*(S)\end{aligned}$$

$$\text{Bel}_*(R_*(S))=\sum_{A\subseteq R_*(S)} p_*(A)=\text{Bel}_{**}(S)$$

As a consequence of previous results,[23] Bel_{**} is also a belief function, whose equivalent random set is the extension of $(R_*(\mathscr{F}),p_*)$. Indeed

$$\begin{aligned}\mathrm{Bel}_{**}(S) &= \sum_{R_*(T)\subseteq R_*(S)} p(T)\\ &= \sum_{A\subseteq R_*(S)} p_*(A)\\ &= \sum_{\omega(A)\subseteq S} p_*(A)\end{aligned}$$

since $R_*(T)\subseteq R_*(S)$ is equivalent to $\omega(R_*(T))\subseteq S$. Bel_{**} is thus a coarsened belief function as much as $\mathrm{Bel}^*{}_*$, but has not been considered by Shafer *et al.*[22]

Now remembering the relationship between the various "rough" inclusions we obtain the following inequalities:

$$\mathrm{Bel}^*{}_*(S)\leqq \mathrm{Bel}(S)\leqq \mathrm{Bel}^{**}(S)$$

$$\mathrm{Bel}_*{}^*(S)\geqq \mathrm{Bel}_{**}(S)\geqq \mathrm{Bel}(S).$$

The first one is also clear from (19–20) which indicate a tight bracketing. However, there is no inequalities between Bel^{**} and Bel_{**}, generally. The interest of the latter inequality lies in the fact that Bel_{**} is a belief function that may act as an upper bound to Bel, so that Bel can be bracketed by two belief functions namely $\mathrm{Bel}^*{}_*$ (from below) and Bel_{**} (from above) that correspond to the upper and lower approximations of $(\mathscr{F}, p)$ and are a generalization of rough fuzzy sets. This result extends Lemmas 11 and 12 of Shafer *et al.*[22] that only consider $\mathrm{Bel}^*{}_*$ as an approximation of Bel on X due to R. Of course dual results can be obtained for the plausibility function since $\mathrm{Bel}(S)=1-\mathrm{Pl}(\bar{S})$.

3. APPROXIMATION OF SETS WITH GRADED SIMILARITY RELATIONS

3.1. Fuzzy Similarity Relations and Partitions

Another extension of rough sets consists in equipping X with a proximity relation, i.e. a fuzzy set R on X^2 such that $\mu_R(x,x)=1$ (reflexivity), $\mu_R(x,y)=\mu_R(y,x)$ (symmetry) and a $_*$-transitivity property of the form

$$\mu_R(x,z)\geqq \mu_R(x,y)_*\mu_R(y,z) \tag{21}$$

for some operation $_*$ satisfying $a_*b\leqq\min(a,b)$. Zadeh[6] has introduced such relations. They generalize equivalence relations in the sense that the core of R, i.e. $\{(x,y)\,|\,\mu_R(x,y)=1\}$ is an equivalence relation. Examples of admissible transitivity axioms are obtained for $*=\min$ (Zadeh's similarity relations), $_*=$ product, $_*=T_m$ ($aT_mb=\max(0,a+b-1)$), see Bezdek and Harris.[25] Particularly $1-\mu_R$ is an ultrametric distance for $_*=\min$, and satisfies the usual triangle inequality for $_*=T_m$. Trillas and Valverde[26,27] have assumed $_*$ is a triangular norm (Schweizer and Sklar[28]), i.e. a non-decreasing semi-group of the unit interval with unity 1, that subsumes the three basic operations. We thus get close to the numerous works devoted to approximation in metric-like spaces, a brief survey of which is in Fariñas del Cerro and Prade.[7]

A simple way to get a fuzzy relation satisfying (21) with $_*=T_m$ is to start with n classical equivalence relations $R_1,\ldots,R_n$ on X and to define (Bezdek and Harris[25])

$$\mu_R(x, y) = \sum_{i=1,n} \alpha_i \mu_{R_i}(x, y) \tag{22}$$

where $\sum \alpha_i = 1$ and $\alpha_i > 0, \forall_i$. When the R_i are nested ($R_1 \subseteq R_2 \ldots \subseteq R_n$), R satisfies (21) with $* = \min$. However, the problem of finding T_m-transitive proximities decomposing into (22) is still unsolved, except for min-transitive ones for which the R_i's are the (nested) level cuts of R, i.e. $\{\{R_\alpha \mid xR_\alpha y \Leftrightarrow \mu_R(x, y) \geqq \alpha\} \mid \alpha \in [0, 1]\}$ (Zadeh[6]).

Another interesting problem is to define the counterpart of equivalence classes for $*$-transitive proximity relations. In fact that is a basic problem in taxonomy. Following Zadeh[6] we can define the fuzzy class $[x]_R$ of elements close to x by

$$\mu_{[x]_R}(y) = \mu_R(x, y) \quad \forall y \in X \tag{23}$$

This definition coincides with the one of usual equivalence classes when R is a non-fuzzy relation. Moreover the existence of $x \neq y$ such that $\mu_R(x, y) = 1$ ensures that the fuzzy sets $[x]_R$ and $[y]_R$ are equal, due to the max-* transitivity relation. Indeed let $z \in X$; then

$$\mu_R(y, z) \geqq \mu_R(y, x) * \mu_R(x, z) = \mu_R(x, z)$$

$$\mu_R(x, z) \geqq \mu_R(x, y) * \mu_R(y, z) = \mu_R(y, z).$$

Hence $\forall z, \mu_R(x, z) = \mu_R(y, z)$. As a consequence the fuzzy equivalence classes $[x]_R$ allow for an extension of rough sets to account for graded indistinguishibility. But the set X/R is now a collection of fuzzy sets that makes a "fuzzy partition" of X. However, note that if the core R_1 of R is a very fine equivalence relation, i.e. such that $x \neq y \Rightarrow [x]_{R_1} \neq [y]_{R_1}$, then $[x]_R \neq [y]_R$ as well so that $X/R = \{[x]_R \mid x \in X\}$ contains as many elements as X.

A more direct way of obtaining a fuzzy coarsening is to start with a family Φ of normal fuzzy sets of X, say $F_1, F_2, \ldots, F_n$, with $n < |X|$ generally. Φ is supposed to cover X sufficiently, i.e.

$$\inf_x \max_{i=1,n} \mu_{F_i}(x) > 0. \tag{24}$$

Moreover, a disjointness property between the F_i's can be requested, e.g.

$$\forall i, j, \sup_x \min(\mu_{F_i}(x), \mu_{F_j}(x)) < 1 \tag{25}$$

This is the weakest possible view of a fuzzy partition; in the literature (e.g. Bezdek[29]) a stronger definition is often adopted i.e.

$$\sum_{i=1,n} \mu_{F_i}(x) = 1, \quad \forall x \in X.$$

However it is not requested here. $F_1, \ldots, F_n$ play the role of fuzzy equivalence classes of a similarity relation, e.g. $\Phi = \{[x]_R \mid x \in X\}$. This situation often occurs when a measurement scale X is shared into a few linguistic categories, e.g. $[0,2]$ meters shared into fuzzy sets which mean "*short*", "*medium-sized*", "*tall*", "*very tall*",.... In that case the term set $\Phi = \{F_1, \ldots, F_n\}$ plays the role of the quotient set X/R. The problem of deriving a fuzzy partition from a fuzzy similarity relation is solved by (23). Indeed X/R satisfies (24) and (25): the fuzzy union of the $[x]_R$ is exactly X: (24) holds and its left hand side is 1. Moreover if $[x]_R \neq [y]_R$ then $\nexists z, \mu_R(x,z) = \mu_R(y,z) = 1$ hence (25) holds too.

The converse problem i.e. given a family of fuzzy sets on X that represents clusters of similar elements, find the underlying relation has been solved by Valverde.[27] Namely given a triangular norm[28] $*$ on $[0,1]$ let $* \!\to$ be a multiple-valued implication that derives from $*$ by residuation, i.e.

$$a * \!\to b = \sup \{x \in [0,1] \mid a * x \leqq b\}$$

Note that $a * \!\to b = 1$ as soon as $a \leqq b$, and $1 * \!\to b = b$. Let $F_1, \ldots, F_n$ be fuzzy sets on X. Then the fuzzy relation R defined by

$$\mu_R(x,y) = \min_{i=1,n} (\max(\mu_{F_i}(x), \mu_{F_i}(y)) * \!\to \min(\mu_{F_i}(x), \mu_{F_i}(y))) \tag{26}$$

is a max-$*$ transitive similarity relation. Moreover if the family of fuzzy sets is $X/R = \{[x]_R \mid x \in X\}$, then (26) applied to this family produces R again.

However if we start with any family $\{F_1, F_2, \ldots, F_n\}$ of fuzzy sets on X, the relation R is defined only up to the choice of the transitivity property (via the choice of $*$), and $X/R \neq \{F_1, F_2, \ldots, F_n\}$ generally. Hence the problem is to select the operation $*$ that produces a fuzzy quotient set X/R that is as close as possible to $\{F_1, F_2, \ldots, F_n\}$. See López de Mántaras and Valverde[30] on these topics.

The application of (26) to the 3 basic kinds of transitive similarity relation gives:

—max-min transitivity

$$\mu_R(x,y) = \min_{i:\mu_{F_i}(x) \neq \mu_{F_i}(y)} \min(\mu_{F_i}(x), \mu_{F_i}(y))$$

$$= 1 \text{ otherwise}$$

—max-product transitivity (Ovchinnikov[31])

$$\mu_R(x,y) = \min_{i=1,n} \min\left(\frac{\mu_{F_i}(x)}{\mu_{F_i}(y)}, \frac{\mu_{F_i}(y)}{\mu_{F_i}(x)}\right)$$

—max-linear transitivity

$$\mu_R(x,y) = \min_{i=1,n} 1 - |\mu_{F_i}(x) - \mu_{F_i}(y)|$$

The basic reason why (26) holds by substituting $([x]_R, x \in X)$ to the F_i's in (26) is that it comes down to solve the relational equation (with symmetric and reflexive relations R) derived from the transitivity property. Namely $R \otimes T \subseteq R$ is equivalent to $T \subseteq R \otimes\!\rightarrow R$ where $\otimes$ is the max-* matrix composition and $\otimes\!\rightarrow$ is the min-*$\rightarrow$matrix composition, after results by Sanchez[32] and Pedrycz.[33]

The transitivity property can be viewed as acknowledging R as a solution to $R \otimes T \subseteq R$. Moreover R is also a solution of $T \otimes R \subseteq R$, so that $T \subseteq (R \otimes\!\rightarrow R)^{-1}$ (where for a fuzzy relation S, $\mu_{S^{-1}}(x, y) = \mu_S(y, x)$). As a consequence

$$T \subseteq T^* = (R \otimes\!\rightarrow R)^{-1} \cap (R \otimes\!\rightarrow R)$$

where $\cap$ translates into min, i.e.

$$\mu_{T^*}(x, y) = \min(\min_z(\mu_R(x, z) \overset{*}{\rightarrow} \mu_R(z, y)), \min_{z'}(\mu_R(y, z') \overset{*}{\rightarrow} \mu_R(z', x)))$$

$$= \min_z(\min(\mu_R(x, z) \overset{*}{\rightarrow} \mu_R(y, z), \mu_R(y, z) \overset{*}{\rightarrow} \mu_R(x, z)))$$

(because, R is symmetric, and due to the properties of min).

$$= \min_z(\max(\mu_R(x, z), \mu_R(y, z)) \overset{*}{\rightarrow} \min(\mu_R(x, z), \mu_R(y, z)))$$

$$\text{since} \quad \min(a \overset{*}{\rightarrow} b, c \overset{*}{\rightarrow} b) = \max(a, c) \overset{*}{\rightarrow} b$$

$$\min(a \overset{*}{\rightarrow} b, a \overset{*}{\rightarrow} c) = a \overset{*}{\rightarrow} \min(b, c)$$

By construction $R \subseteq T^*$. Letting $z = x$ in the expression of μ_{T^*}, leads to $T^* \subseteq R$, due to reflexivity and symmetry of R. (26) is completely justified by the identity $T^* = R$.

3.2. *Fuzzy Rough Sets*

Given a fuzzy partition Φ on X, counterparts of (1–2) and (5–6) in this setting allow a description of any fuzzy set F by means of the term set Φ, under the form of an upper and a lower approximation $\Phi^*(F)$ and $\Phi_*(F)$ as follows

$$M_i \triangleq \mu_{\Phi^*(F)}(F_i) = \sup_x \min(\mu_{F_i}(x), \mu_F(x)) \tag{27}$$

$$m_i \triangleq \mu_{\Phi_*(F)}(F_i) = \inf_x \max(1 - \mu_{F_i}(x), \mu_F(x)) \tag{28}$$

Clearly, these equations also generalize the basic definitions of C-calculus. (27) was first proposed by Willaeys and Malvache[34] in order to define a fuzzy set on a referential of fuzzy sets. (27) and (28) are nothing but the degrees of possibility and necessity of the fuzzy event F, in the sense of Zadeh.[12,18] M_i being the degree of possible membership of F_i in F, and m_i the corresponding degree of certain membership. This definition makes sense even when F is not fuzzy. The pair $(\Phi_*(F), \Phi^*(F))$ can be called a fuzzy rough set.

Assume now that Φ derives from a fuzzy relation R via (23). The knowledge of R is anyway sufficient to define the "extension" of fuzzy rough sets in the sense of (3–4). Given a subset S of X, (7) and (8) generalize into

$$\forall x \in X, \mu_{\omega(R^*(S))}(x) = \sup \{\mu_R(x,y) \mid y \in S\} \tag{29}$$

$$\forall x \in X, \mu_{\omega(R_*(S))}(x) = \inf \{1 - \mu_R(x,y) \mid y \notin S\} \tag{30}$$

Moreover (29–30) in turn generalize in accordance with possibility theory into (Fariñas del Cerro and Prade[7])

$$\forall x \in X, \mu_{\omega(R^*(F))}(x) = \sup_y \min(\mu_F(y), \mu_R(x,y)) \tag{31}$$

$$\forall x \in X, \mu_{\omega(R_*(F))}(x) = \inf_y \max(\mu_F(y), 1 - \mu_R(x,y)) \tag{32}$$

which define upper and lower approximations of a fuzzy set through a similarity relation. The equivalence between (27–28) and (31–32), when $\Phi = X/R$, is almost self-evident, since the index i in (27–28) is just replaced by the element x in (31–32), based on (23); moreover, if x and x' generate the same fuzzy equivalence class, their degrees of membership in (31–32) coincide. Besides, (31) and (32) represent the distortion of a fuzzy set F due to the indiscernibility relation on X, as if looking at F with blurring glasses.

Nakamura[35] has also suggested definitions of fuzzy rough sets. He defines a family of upper approximations of F, $\{R^*_\alpha(F), \alpha \in (0,1]\}$ using (7) on each α-cut R_α, much in the spirit of Orlowski[36]'s fuzzy optimization model. He also introduces a family of lower approximations in the same spirit, using (8), that is: $\{R^*_\alpha(F), \alpha \in (0,1]\}$. Note that this view makes sense only when R_α is an equivalence relation for all α in the unit interval, that is, if R is max-min transitive. Then R defines a consistent set $\{X/R_\alpha \mid \alpha \in (0,1]\}$ of more or less coarsened versions of X, and Nakamura's approach leads to as many rough approximations of a fuzzy set F on these coarser referential sets X/R_α. The higher α, the finer the partition. (31) and (32) define a fuzzy rough set that can be viewed as the weighted union of the $R^*_\alpha(F)$'s and the weighted intersection[37] of the R^*_α's respectively, that is:

$$\mu_{R^*(F)}(F_i) = \sup_\alpha \min(\alpha, \mu_{R^*_\alpha(F)}(F_i)) \tag{33}$$

$$\mu_{R_*(F)}(F_i) = \inf_\alpha \max(1 - \alpha, \mu_{R^*_\alpha(F)}(F_i)) \tag{34}$$

Proof It is easy to check that if $\Pi(F \mid G)$ denotes the possibility of the fuzzy event F based on possibility distribution μ_G then using α-cuts[38]

$$\Pi(F \mid G) = \sup_x \min(\mu_F(x), \mu_G(x)) = \sup_{\alpha,\beta} \min(\alpha, \beta, \Pi(F_\beta \mid G_\alpha))$$

where $\Pi(F_\beta \mid G_\alpha) = 1$ if $F_\beta \cap G_\alpha \neq \varnothing$, and 0 otherwise. As a consequence $\Pi(F \mid G) = \sup_\alpha \min(\alpha, \Pi(F \mid G_\alpha))$. This is formula (33) for $\mu_G = \pi_i$, the membership function of F_i, since $\mu_{R^*(F)}(F_i) = \Pi_i(F)$, as in (7).

Moreover $N(F \mid G) = 1 - \Pi(\bar{F} \mid G) = 1 - \sup_\alpha \min(\alpha, \Pi(\bar{F} \mid G_\alpha))$

$$= \inf_{\alpha} \max(1-\alpha, 1-\Pi(\bar{F}|G_\alpha)) = \inf_{\alpha} \max(1-\alpha, N(F|G_\alpha)).$$

This is (34) noticing that $\mu_{R_*(F)}(F_i) = N_i(F)$. Q.E.D.

Moreover, (33) and (34) represents weighted medians of the sets $\{\mu_{R^*_\alpha(F)}(F_i) | \alpha \in [0,1]\}$ and $\{\mu_{R^*_\alpha(F)}(F_i) | \alpha \in [0,1]\}$ in the sense of Sugeno's integral.[39]

N.B.: (27–28) suggest a new approach to linguistic approximation. Assume $X/R = \{F_1, F_2, \ldots, F_n\}$ is a term set that partition X into linguistically meaningful fuzzy subsets. The problem of linguistic approximation is one of finding the best term that may qualify a given unnamed fuzzy set F. This problem is commonly encountered in fuzzy set-based software that accept linguistic inputs, model them by fuzzy sets, and must output responses in a linguistic format again. Then the fuzzy sets computed by the program must be named by means of some user-oriented vocabulary. See for instance Novak,[40]... If I_* and I^* denote the set of indices such that

$$m_i = \mu_{R_*(F)}(F_i) > 0 \quad \text{and} \quad M_i = \mu_{R^*(F)}(F_i) > 0$$

respectively, F can be bracketed by ORing the F_i's in I_* and in I^* (which contains I_*). The M_i's and m_i's, may be used to define linguistic modifiers for the selected terms in the term set. Namely, F means "$L_{i1} F_{i1}$ or $L_{i2} F_{i1}$ or ... $L_{ip} F_{ip}$" where p is the cardinality of I^*, and L_i is a modifier of F_i. Roughly speaking, the following modifiers could be chosen:

if $M_i \leqq 0.5$	L_i means "not very possibly"
if $m_i = 0, M_i \geqq 0.5$	L_i means "possibly"
if $m_i \geqq 0.5$	L_i means "rather certainly"
if $m_i = 1$	L_i means "certainly" (and can be omitted)

This approach to linguistic approximation is quite different from the one of Bonissone[9] and followers, based on minimizing a distance between F and the fuzzy sets produced from the F_i's by means of unary modifiers and logical connectives. Here we view linguistic approximation as a problem of rough classification in the sense of Pawlak *et al.*[41] Our procedure looks more robust than the ones based on a nearest-neighbour classification process. But it cannot precisely name fuzzy sets included in the F_i's, by construction.

3.3. *Properties of Fuzzy Rough Sets*

Basic properties of rough sets can be extended to fuzzy rough sets. Using classical pointwise definitions of union, intersection and complementation, by maximum, minimum, and 1-. respectively, and the usual definition of fuzzy set inclusion, the following properties hold on the term set $\Phi = X/R$

$$\Phi_*(F \cup G) \supseteq \Phi_*(F) \cup \Phi_*(G); \Phi^*(F \cup G) = \Phi^*(F) \cup \Phi^*(G) \tag{35}$$

$$\Phi^*(F \cap G) \subseteq \Phi^*(F) \cap \Phi^*(G); \Phi_*(F \cap G) = \Phi_*(F) \cap \Phi_*(G) \tag{36}$$

$$\Phi^*(\bar{F}) = \overline{\Phi_*(F)} \tag{37}$$

where inclusion is in the sense of fuzzy set theory. They are another way of stating basic relations between $N(F)$, $N(G)$, $N(F \cup G)$, $N(F \cap G)$ on the one hand, and $\Pi(F)$, $\Pi(G)$, $\Pi(F \cup G)$, $\Pi(F \cap G)$, on the other hand, in possibility theory.

When the fuzzy rough set derives from a fuzzy relation R, then, we can check the following properties[7]

$$\omega(R_*(F)) \subseteq F \subseteq \omega(R^*(F)) \tag{38}$$

$$R_{**}(F) \subseteq R_*(F); R^{**}(F) \supseteq R^*(F) \tag{39}$$

where $R_{**}(F)$ is short for $R_*(\omega(R_*(F)); R^{**}(F) = R^*(\omega(R^*(F))$.To check (39) it is enough to notice that

$$\begin{aligned}\sup_y \min(\mu_F(y), \mu_R(x,y)) &\geqq \max\left(\mu_F(x), \sup_{y \neq x} \min(\mu_R(x,y), \mu_F(y))\right)\\ &\geqq \mu_F(x)\end{aligned}$$

since $\mu_R(x,x) = 1$ (reflexivity of R). This solves the second inclusion in (38). The first one is done similarly or using (37). (39) is due to (38) changing F into $R_*(F)$ and $R^*(F)$ respectively. Moreover it is always true that $R^{**}(F) = (R \circ R)^*(F)$, where $\circ$ denote the sup-min composition of fuzzy relations, since

$$\begin{aligned}&\sup_y \min\left(\sup_z \min(\mu_F(z), \mu_R(z,y)), \mu_R(y,x)\right)\\ &\quad = \sup_z \min\left(\mu_F(z), \sup_y \min(\mu_R(z,y), \mu_R(y,x))\right)\\ &\quad = \sup_z \min(\mu_F(z), \mu_{R \circ R}(z,x)).\end{aligned}$$

Note that if R is max-min transitive, then $R \circ R = R$ and the inequalities in (39) become equalities, as in standard rough set theory.[2] However, generally,

$$R^*(\omega(R_*(F)) \supset \omega(R_*(F))$$

$$R_*(\omega(R^*(F)) \subset \omega(R^*(F))$$

contrary to what happens with crisp rough sets where equality holds. More specifically, if R is max-min transitive,[7]

$$\forall x, \mu_{R_*(F)}(x) \leqq \mu_{R^*(\omega(R_*(F)))}(x) \leqq \max(\mu_{R_*(F)}(x), 0.5)$$

For general kinds of fuzzy partitions with normalized fuzzy sets, we also have $\Phi^*(\Phi^*(F)) \supseteq \Phi^*(F)$, $\Phi_*(\Phi_*(F)) \subseteq \Phi_*(F)$, viewing $\Phi^*(F)$ and $\Phi_*(F)$ as weighted unions[37] of F_i's, with weights M_i and m_i respectively i.e. $\forall x \in X$, $\mu_{\Phi^*(F)}(x) = \max_{i=1,n} \min(M_i, \mu_{F_i}(x))$ for the upper approximation, and similarly for $\mu_{\Phi_*(F)}$ changing M_i into m_i.

6. CONCLUSION

This paper has shown that the idea of a rough set can be combined with fuzzy sets in a fruitful way. It enables several independent approaches to approximation models to be unified. Some lines of research have been mentioned in the previous sections. Let us add two other ones. First it might be useful to generalize (27) and (28) changing min into a general intersection operation, and using another inclusion index in (28). Indeed, it may be checked that the upper and lower approximation of F_i in (27) and (28) is different from F_i, generally; this is because in (28) the inclusion index is not compatible with Zadeh's inclusion based on the inequality between membership functions. Second, Pawlak[10] as well as Shafer *et al.*[22] and Pearl and Verma[42] have studied concepts of independence and redundancy between partitions, each with different terminologies. Especially Pawlak's notion of independence is weaker than the one of Shafer *et al.*[22]. It might be useful to pursue the unification work along this line, and generalize it with fuzzy sets.

REFERENCES

1. L. A. Zadeh, "Fuzzy sets." *Information & Control*, **8**, 1965, pp. 338–353.
2. Z. Pawlak, "Rough sets." *International Journal of Computer and Information Science*, **11**, 1982, pp. 341–356.
3. Z. Pawlak, "Rough sets and fuzzy sets." *Fuzzy Sets & Systems*, **17**, 1985, pp. 99–102.
4. E. R. Caianiello, "A calculus for hierarchical systems." *Proc. 1st Inter. Conf. on Pattern Recognition*, Washington, D.C., 1973, pp. 1–5.
5. G. Shafer, *A Mathematical Theory of Evidence*. Princeton Univ. Press, Princeton, N.J., 1976.
6. L. A. Zadeh, "Similarity relations and fuzzy orderings." *Information Sciences*, **3**, 1971, pp. 177–200.
7. L. Fariñas del Cerro and H. Prade, "Rough sets, twofold fuzzy sets and modal logic—Fuzziness in indiscernibility and partial information." In: *The Mathematics of Fuzzy Systems*, edited by A. Di Nola and A. G. S. Ventre, Verlag TÜV Rheinland, Köln, 1986, pp. 103–120.
8. L. A. Zadeh. "Quantitative fuzzy semantics." *Information Sciences*, **3**, 1971, pp. 159–176.
9. P. P. Bonissone, "A pattern recognition approach to the problem of linguistic approximation in system analysis." *Proc. IEEE Inter. Conf. on Cybernetics and Society*, Denver, 1979, pp. 793–798.
10. Z. Pawlak, "Rough classification." *Int. J. Man-Machine Stud.*, **20**, 1984, pp. 469–485.
11. V. Tahani, "A conceptual framework for fuzzy query processing—A step towards very intelligent database systems." *Information Processing & Management*, **13**, 1977, pp. 289–303.
12. L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, **1**, 1978, pp. 3–28.
13. H. Prade and C. Testemale, "Generalizing database relational algebra for the treatment of uncertain/imprecise information and vague queries." *Information Sciences*, **34**, 1984, pp. 27–42.
14. D. Dubois and H. Prade, "Twofold fuzzy sets and rough sets—Some issues in knowledge representation," *Fuzzy Sets & Systems*, **23**, 1987, pp. 3–18.
15. S. M. K. Wong and W. Ziarko, "A probabilistic model of approximate classification and decision rules with uncertainties and inductive learning." *Technical Report CS-85-23*, University of Regina, Saskatchewan, 1985.
16. A. Pawlak, S. M. K. Wong and W. Ziarko, "Rough sets: probabilistic versus deterministic approach." *Int. J. of Man-Machine Studies*, **29**, 1988, pp. 81–95.
17. M. Wygralak, "Rough sets and fuzzy sets: some remarks on interrelations." *Fuzzy Sets & Systems*, **29**, 1989, pp. 241–243.

18. D. Dubois and H. Prade (with the collaboration of H. Farreny, R. Martin-Clouaire and C. Testemale) (1988) *Possibility Theory: An Approach to Computerized Processing of Uncertainty.* Plenum Press, New York.
19. E. R. Caianiello, "C-calculus: an overview." In: *Topics in the General Theory of Structures,* edited by E. R. Caianiello and M. A. Aizerman, D. Reidel Dordrecht, 1987, pp. 163–173.
20. E. R. Caianiello and A. G. S. Ventre, "C-calculus and fuzzy sets." *Proc. 1st Napoli Meeting on the Mathematics of Fuzzy Systems,* Università degli Studi di Napoli, 1984, pp. 29–33.
21. E. R. Caianiello and A. G. S. Ventre, "A model for C-calculus." *Int. J. General Syst.,* **11,** 1985, pp. 153–161.
22. G. Shafer, P. Shennoy and K. Mellouli, "Qualitative Markov trees." *Int. J. Approximate Reasoning,* **1,** 1987, pp. 349–400.
23. D. Dubois and H. Prade, "Evidence measures based on fuzzy information." *Automatica,* **21,** 1985, pp. 547–562.
24. A. P. Dempster, "Upper and lower probabilities induced by a multiple-valued mapping." *Annnals of Mathematical Statistics,* **38,** 1967, pp. 325–339.
25. J. C. Bezdek and J. D. Harris, "Fuzzy partitions and relations—An axiomatic basis for clustering." *Fuzzy Sets & Systems,* **1,** 1978, pp. 112–127.
26. E. Trillas and L. Valverde, "An investigation into indistinguishability operators." In: *Aspects of Vagueness,* edited by H. J. Skala, S. Termini and E. Trillas, D. Reidel Dordrecht, 1984, pp. 231–256.
27. L. Valverde, "On the structure of F-indistinguishability operators." *Fuzzy Sets & Systems,* **17,** 1985, pp. 313–328.
28. B. Schweizer and A. Sklar, *Probabilistic Metric Spaces.* North-Holland, Amsterdam, 1983.
29. J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* Plenum Press, New York, 1981.
30. R. López de Mántaras and L. Valverde, "New results in fuzzy clustering based on the concept of indistinguishability relation." *IEEE Trans. on Pattern Analysis and Machine Intelligence,* **10,** 1988, pp. 754–757.
31. S. V. Ovchinnikov, "On fuzzy relational systems." *Proc. 2d World Conf. on Mathematics at the Service of Man,* Las Palmas, Spain, 1982, pp. 566–569.
32. E. Sanchez, "Resolution of composite fuzzy relation equations." *Information and Control,* **30,** 1976, pp. 38–48.
33. W. Pedrycz, "On generalized fuzzy relational equations and their applications." *J. of Mathematical Analysis and Applications,* **107,** 1985, pp. 520–536.
34. D. Willaeys and N. Malvache. "The use of fuzzy sets for the treatment of fuzzy information by computer." *Fuzzy Sets and Systems,* **5,** 1981, pp. 323–328.
35. A. Nakamura, "Fuzzy rough sets." *Note on Multiple-Valued Logic in Japan,* **9**(8), 1988, pp. 1–8.
36. S. A. Orlowski, "On programming with fuzzy constraint set." *Kybernetes,* **6,** 1977, pp. 197–201.
37. D. Dubois and H. Prade, "Weighted minimum and maximum in fuzzy set theory." *Information Sciences,* **39,** 1986, pp. 205–210.
38. H. Prade, "Modal semantics and fuzzy set theory." In: *Fuzzy Set and Possibility Theory: Recent Developments,* edited by R. R. Yager, Pergamon Press, Oxford, U.K., 1982, pp. 232–246.
39. M. Sugeno, "Fuzzy measures and fuzzy integrals: a survey." In: *Fuzzy Automated and Decision Processes,* edited by M. M. Gupta, G. N. Saridis and B. R. Gaines, North-Holland, Amsterdam, 1977, pp. 89–102.
40. V. Novak. "Automatic generation of verbal comments on results of mathematical modelling. In: *Approximate Reasoning in Intelligent Systems, Decision and Control,* edited by E. Sanchez and L. A. Zadeh, Pergamon Press, Oxford, U.K., 1987, pp. 55–68.
41. Z. Pawlak, K. Slowinski and R. Slowinski, "Rough classification of patients after highly selective vagotomy for duodenal ulcer." *Int. J. Man-Machine Stud.,* **24,** 1986, pp. 413–433.
42. J. Pearl and T. Verma, "The logic of representing dependencies by directed graphs." *Proc. 6th AAAI National Conf. on Artificial Intelligence,* Seattle, 1987, pp. 374–379.

Didier Dubois (Engineer degree (1975) and Doctor-Engineer degree (1977), Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, Toulouse, France; "Doctorat d'Etat", University of Grenoble, 1983; "Habilitation à Diriger des Recherches", University of Toulouse, 1986) is a full-time researcher at the National Center for Scientific Research (C.N.R.S.). From 1980 till 1983, he worked as a research engineer at the Centre d'Etudes et de Recherches de Toulouse, in the Production Research area. He co-authored with Henri Prade two books on fuzzy sets and possibility theory, in 1980 and 1985, respectively, and has co-edited a volume on *Non-Standard Logics* and several special issues of scientific journals. He belongs to Editorial Boards of several journals such as *Fuzzy Sets and Systems*, *International Journal of Approximate Reasoning*, the French *Revue d'Intelligence Artificielle*, and more recently the *International Journal of General Systems* and the *ORSA Journal of Computing*. His main topics of interest are the modelling of imprecision and uncertainty, the representation of knowledge and approximate reasoning for expert systems, operations research and decision analysis. He is an author or co-author of numerous publications, especially in the field of fuzzy sets and their applications to operations research and artificial intelligence.

Henri Prade (Engineer degree (1975) and Doctor-Engineer degree (1977), Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, Toulouse, France; "Doctorat d'Etat", University of Toulouse, 1982; "Habilitation à Diriger des Recherches", University of Toulouse, 1986) is a full-time researcher at the National Center for Scientific Research (C.N.R.S.). He is an author (with Dr. Didier Dubois) of two books: *Fuzzy Sets and Systems: Theory and Applications* (Academic Press, New York, 1980) and *Possibility Theory* (Masson, Paris, 1985; English translation: Plenum Press, New York, 1988), of over one hundred technical papers, and a co-editor of two books as well as several special issues of international journals. He is a member of Editorial Boards of *Fuzzy Sets and Systems*, *Inter. J. of Intelligent Systems*, *Inter. J. of Approximate Reasoning*, *Inter. J. of Expert Systems: Research and Applications*, *Inter. J. Computers and Artificial Intelligence*, and the French *Revue d'Intelligence Artificielle*. His current research interests are in fuzzy set and possibility theory, approximate reasoning and non-classical logics, artificial intelligence and expert systems, incomplete and uncertain information data base management systems, and operations research.