Rough Mereology as a Link Between Rough and Fuzzy Set Theories. A Survey

Lech Polkowski[1,2]

[1] Polish–Japanese Institute of Information Technology
Koszykowa 86, 02 008 Warsaw, Poland
[2] Dept. Math.Computer Science, Univ. Warmia and Mazury
Zolnierska 14 A, 10 561 Olsztyn, Poland
Lech.Polkowski@pjwstk.edu.pl

Abstract. In this paper, we discuss rough inclusions defined in Rough Mereology – a paradigm for approximate reasoning introduced by Polkowski and Skowron – as a basis for common models for rough as well as fuzzy set theories. We justify the point of view that tolerance (or, similarity) is the motif common to both theories. To this end, we demonstrate in Sect. 6 that rough inclusions (which represent a hierarchy of tolerance relations) induce rough set theoretic approximations as well as partitions and equivalence relations in the sense of fuzzy set theory. Before that, we include an account of mereological theory in Sect. 3. We also discuss granulation mechanisms based on rough inclusions with an outline of applications to rough–fuzzy–neurocomputing and computing with words in Sects. 4 and 5.

Keywords: rough set theory, fuzzy set theory, rough mereology, rough inclusions, granular calculus, rough-fuzzy-neurocomputing, computing with words

1 Introduction

We begin with a concise review of both theories.

1.1 Rough Sets: Basic Ideas

Rough Set Theory begins with the idea, [16], [17], of an *approximation space*, understood as a universe U together with a family $\mathcal{R}$ of equivalence relations on U (knowledge base). Given a sub–family $\mathcal{S} \subseteq \mathcal{R}$, the equivalence relation $S = \bigcap \mathcal{S}$ induces a partition P_S of U into equivalence classes $[x]_S$ of the relation S.

In terms of P_S, *concept approximation* is possible; a *concept relative to* U (or, shortly, a U*–concept*) is a subset $X \subseteq U$. There are two cases.

A U–concept X is S*–exact* in case,

$$X = \bigcup \{[x]_S : [x]_S \subseteq X\}, \tag{1}$$

holds. We will speak in this case also of an S*–exact set*.

Otherwise, X is said to be an S*–rough* U-concept (or, set).

J.F. Peters et al. (Eds.): Transactions on Rough Sets II, LNCS 3135, pp. 253–277, 2004.

In the second case, the idea of an *approximation* comes useful [16]. Two S–exact sets, approximating X from below and from above, are the *lower S–approximation*,

$$\underline{S}X = \bigcup\{[x]_S : [x]_S \subseteq X\}, \tag{2}$$

and the *upper S–approximation*,

$$\overline{S}X = \bigcup\{[x]_S : [x]_S \cap X \neq \emptyset\}. \tag{3}$$

Then, clearly,

1. $\underline{S}X \subseteq X \subseteq \overline{S}X$.
2. $\underline{S}X$ (respectively, $\overline{S}X$) is the largest (respectively, the smallest) S–exact set contained in (respectively, containing) X.

Sets (U–concepts) with identical approximations may be identified; consider an equivalence relation $\approx_S$ defined as follows [18]:

$$X \approx_S Y \Leftrightarrow \underline{S}X = \underline{S}Y \wedge \overline{S}X = \overline{S}Y. \tag{4}$$

This is clearly an equivalence relation; let $Concepts_S$ denote the set of these classes. Then for $x, y \in Concepts_S$, we have

$$x = y \Leftrightarrow \forall u \in U.\phi_u(x) = \phi_u(y) \wedge \psi_u(x) = \psi_u(y), \tag{5}$$

where for $x = [X]_{\approx_S}$, $\phi_u(x) = 1$ in case $[u]_S \subseteq X$, otherwise $\phi_u(x) = 0$; similarly, $\psi_u(x) = 1$ when $[u]_S \cap X \neq \emptyset$, otherwise $\psi_u(x) = 0$.

The formula (5) witnesses *Leibnizian indiscernibility* in $Concept_S$: entities are distinct if and only if they are discerned by at least one of available functionals (in our case, ϕ_u, ψ_u). The idea of **indiscernibility** is one of the most fundamental in Rough Set Theory [16].

Other fundamental notions are derived from the observation on complexity of the generation of S: one may ask whether there is some $\mathcal{T} \subset \mathcal{S}$ such that

$$S = \bigcap \mathcal{T}. \tag{6}$$

In case the answer is positive, one may search for a minimal with respect to inclusion subset $\mathcal{T} \subseteq \mathcal{S}$ satisfying (6). Such a subset is said to be an S–**reduct**. Let us observe that $\mathcal{T} \subseteq \mathcal{S}$ is an S–reduct if and only if for each $R \in \mathcal{T}$ we have,

$$\bigcap(\mathcal{T} \setminus \{R\}) \neq \bigcap \mathcal{T}. \tag{7}$$

In this case we say that $\mathcal{U}$ is **independent**; this notion falls under a general independence scheme [14].

These ideas come fore most visibly in case of knowledge representation in the form of *information systems* [16]. An *information system* is a universe U along with a set A of *attributes* each element a of which is a mapping $a : U \to V_a$ from

U into an *a–value set* V_a. Clearly, each attribute $a \in A$ does induce a relation of *a–indiscernibility* $IND(a)$ defined as follows,

$$xIND(a)y \Leftrightarrow a(x) = a(y). \tag{8}$$

The family $\{Ind(a) : a \in A\}$ is a knowledge base and for each $B \subseteq A$, the relation $IND(B) = \bigcap\{IND(a) : a \in B\}$ is defined, inducing the lower and the upper approximations $\underline{B}X, \overline{B}X$. Notions of a B–reduct, and B–independence are defined as in general case [16].

From now on, we will work with the reduced universe $U/IND(A)$ without mentioning this fact explicitly; thus, the identity $=_U$ on objects – classes representatives – will denote the equality of indiscernibility classes.

1.2 Fuzzy Sets: Basic Notions

A starting point for Fuzzy Set Theory is that of a *fuzzy set* [26]. Fuzzy sets come as a generalization of the usual mathematical idea of a set: given a universe U, a set X in U may be expressed by means of its *characteristic function* χ_X: $\chi_X(u) = 1$ in case $u \in X$, $\chi_X(u) = 0$, otherwise.

A fuzzy set X is defined by allowing χ_X to take values in the interval $[0, 1]$. Thus, $\chi_X(u) \in [0, 1]$ is a measure of *degree to which u is in X*. Once fuzzy sets are defined, one may define a *fuzzy algebra of sets* by defining operators responsible for the union, the intersection, and the complement in the realm of fuzzy sets. Usually, those are defined by selecting a *t–norm*, a *t–co–norm*, and a *negation* functions where a t–norm $T(x, y)$ is a function allowing the representation (see [13], cf. [21] Ch. 14),

$$T(x, y) = g(f(x) + f(y)), \tag{9}$$

where the function $f : [0, 1] \to [0, +\infty)$ in (9) is continuous decreasing on $[0, 1]$ and g is the pseudo–inverse to f (i.e. $g(u) = 0$ in case $u \in [0, f(1)]$, $g(u) = f^{-1}(u)$ in case $u \in [f(1), f(0)]$, and $g(u) = 1$ in case $u \in [f(0), +\infty)$).

A t–conorm C is induced by a t–norm T via the formula $C(x, y) = 1 - T(1 - x, 1 - y)$. A negation $n : [0, 1] \to [0, 1]$ is a continuous decreasing function such that $n(n(x)) = x$. An important example of a t–norm is the *Łukasiewicz product*,

$$\otimes(x, y) = max\{0, x + y - 1\}; \tag{10}$$

we recall also the *Menger product*,

$$Prod(x, y) = x \cdot y. \tag{11}$$

1.3 Rough Membership Functions

In a rough universe represented by an information system (U, A), it is possible to introduce, as shown by Pawlak and Skowron [19], a parameterized family of functions measuring for a given object $u \in U$ and a set $X \subseteq U$ degree to which

u belongs in X, a parameter being a subset $B \subseteq A$ of attributes. In this case, the *rough membership function* $m_{B,X}$ on U is defined as,

$$m_{B,X}(u) = \frac{|X \cap [u]_B|}{|[u]_B|}. \tag{12}$$

It may be noticed that [19],

1. In case X is a B–exact set, $m_{B,X}$ is the characteristic function of X, i.e., X is perceived as a crisp set;
2. $m_{B,X}$ is a piece–wise constant function, constant on classes of $IND(B)$.

From $m_{B,X}$, rough set approximations are reconstructed as follows,

$$\underline{B}X = \{u \in U : m_{B,X}(u) = 1\}, \tag{13}$$

$$\overline{B}X = \{u \in U : m_{B,X}(u) > 0. \tag{14}$$

Rough membership functions are prototypes of *rough inclusions* defined in the sequel which measure degrees to which one object (concept) is contained in another, and containment is expressed in terms of relations of being a part.

Let us observe a contextual character of rough membership functions: they are defined relative to sets of attributes, hence, they represent information content of those sets. A similar character will be featured in rough inclusions defined in the framework of information systems.

2 Rough Mereology

We introduce a reasoning mechanism based on Rough Mereology [22], whose distinct facets correspond to rough respectively fuzzy approaches to reasoning allowing on one hand the introduction of rough set–theoretic approximations and on the other hand inducing fuzzy set–theoretic notions of a partition as well as of an equivalence relation. We will present the main ideas of this approach.

2.1 Mereology

Mereological theories, contrary to naive or formalized set theories that are based on the notion of being an element of a set, are based on the idea that relations among object should be based on their containment properties (e.g., we cannot say "the circle is an element of the closed disk it is boundary of" but we can say "the circle is the part of the closed disk"). Mereology theory proposed by Stanisław Leśniewski proposes the notion of a *part* as the primitive one, whereas the mereological theory outlined by Alfred North Whitehead [24] and developed by Leonard, Goodman, and Clarke, see [4], among others, begins with the notion of *connection.*

We work here in the formalism proposed by Leśniewski, and thus the basic notion is that of a *part* relation on a universe U with identity $=$, [11]; in symbols:

$x\pi y$, that reads x *is a part of* y. A part relation π satisfies by definition the following conditions,

$$x\pi y \wedge y\pi z \Rightarrow x\pi z, \tag{15}$$

$$\neg(x\pi x). \tag{16}$$

Thus, the part relation is transitive and non–reflexive (i.e., it is a pre–order on U), so it can be interpreted as expressing the idea of a *proper part*.

The idea of an improper part, i.e., possibly of the whole object, is expressed by the notion of an *element*, that may not be confused with the notion of an element in naive set theory (sometimes, the term "ingredient" is used in place of that of "element", such was original usage of Leśniewski).

The notion of an *element* relation, el_π, induced by the relation π is the following,

$$x\ el_\pi\ y \Leftrightarrow x\pi y \vee x = y. \tag{17}$$

Thus, $x\ el_\pi\ y \wedge y\ el_\pi\ x$ is equivalent to $x = y$. By (15), (16), the relation el_π is an ordering on the universe U. Moreover,

$$xel_\pi x \tag{18}$$

for each x in U, a striking difference with the usage of the notion of an element in naive set theory.

The fundamental feature of mereology of Leśniewski is that it is concerned with collective classes, i.e., distributive classes (names, concepts) are made into objects (collective classes).

In order to make a non–empty concept $M \subseteq U$ into a collective object, the *class* operator Cls is used [11].

The definition of the class of M, in symbols $Cls(M)$, is as follows,

$$x \in M \Rightarrow x\ el_\pi\ Cls(M), \tag{19}$$

$$x\ el_\pi\ Cls(M) \Rightarrow \exists y, z.y\ el_\pi\ x \wedge y\ el_\pi\ z \wedge z \in M. \tag{20}$$

Condition (19) includes all members of M into $Cls(M)$ as elements; (20) requires each element of $Cls(M)$ to have an element in common with a member of M (compare this with the definition of the union of a family of sets to see the analogy that becomes clear, when the relation of being a (mereological) element is interpreted as that of being a subset of a set, and M is a family of subsets in a given universe).

One requires also $Cls(M)$ to be unique, [11]. From this demand it follows that, [11],

$$[\forall y.(y\ el_\pi\ x \Rightarrow \exists w.w\ el_\pi\ y \wedge w\ el_\pi\ z)] \Rightarrow x\ el_\pi\ z. \tag{21}$$

The rule (21) is useful for recognizing that $x\ el_\pi\ z$, and it will be used in our arguments.

2.2 Rough Mereology

Given a mereological universe (U, π), with the induced relation el_π of being an element, we introduce, cf. [22], on $U \times U \times [0,1]$ a relation $\mu(x, y, r)$, read *x is a part of y to degree at least r*. We will write rather more suggestively $x\mu_r y$, calling μ a *rough inclusion* (it is the generic name introduced in [22]).

We introduce at this moment an element of ontology (we do not enter here any discussion on what is ontology; here, by ontology we mean simply a system of names we use for naming various types of objects): we will regard U as a name for atomic objects, and the formula $x\varepsilon U$ (read: *x is U*) will mean that an object x is an atomic object.

In case of an information system (U, A), atomic objects will be objects $x \in U$. The symbol G_U will name individual objects constructed as classes of atomic objects, which will be denoted with the formula $g\varepsilon G_U$, i.e., $g\varepsilon G_U$ if and only if there exists a non–empty property Φ with $g = Cls(\Phi|U)$, $\Phi|U$ meaning Φ restricted to atomic objects, $\Phi|U$ a non–empty property of atomic objects.

We require μ to satisfy the following conditions, see [22],

$$x\mu_1 x, \tag{22}$$

$$x\mu_1 y \Leftrightarrow x\ el_\pi\ y, \tag{23}$$

$$x\mu_1 y \Rightarrow \forall z, r.(z\mu_r x \Rightarrow z\mu_r y), \tag{24}$$

$$x\mu_r y \wedge s < r \Rightarrow x\mu_s y. \tag{25}$$

Informally, (23) ties the rough inclusion to the mereological underlying universe , (24) does express monotonicity (a bigger entity cuts a bigger part of everything), (25) says that a part to degree r is a part to any lesser degree.

We may observe that (22) is a consequence to (23) due to (18).

2.3 More Ontology for Information Systems

We would like to set here some scheme for relating ontology of an information system outlined above to ontology of the associated mereological universe (again, we use term “ontology” here as a name for a system of notions only).

Given an information system (U, A), we introduce the material identity $=_U$ on U (understood as the identity of objects, e.g., identity of patients, signals, etc., witnessed by their names, times of receiving, etc.), and we let for $x, y\varepsilon U$,

$$x\mu_1 y \Leftrightarrow x =_U y. \tag{26}$$

Thus the notion of an *element*, corresponding to the rough inclusion μ by means of (23) coincides with $=_U$ on atomic objects; we will denote with the symbol el_U this notion of an element. We extend el_U to pairs $x \in g$, where $x\varepsilon U, g\varepsilon G_U$. The corresponding notion of a class, Cls_U coincides with the notion of a set. The mereological notion of a subset,

$$xsub_U y \Leftrightarrow \forall z.(zel_U x \Rightarrow zel_U y), \tag{27}$$

coincides with set inclusion $\subseteq$. Thus the mereological complement x^c (defined as the class of all objects that do not have any object as a common part with x) coincides with the set complement $U \setminus x$.

For each $x \varepsilon U$, the set,

$$Inf(x) = \{(a, a(x)) : a \in A\}, \tag{28}$$

will be defined, called the *information set of* x; the universe,

$$INF(U) = \{Inf(x) : x \varepsilon U\}, \tag{29}$$

will be called the *information universe* of the information system (U, A). We let

$$INF : U \rightarrow INF(U), INF(x) = Inf(x) \text{ for } x \varepsilon U. \tag{30}$$

Given a mereological structure on $INF(U)$, we denote with el_{INF}, Cls_{INF}, $=_{INF}$, respectively, the relation of being an element, a class, and the identity on $INF(U)$.

For $g \varepsilon G_U$, we let,

$$INF(g) =_{INF} Cls_{INF}(Inf(y) : y \varepsilon U, y el_U g). \tag{31}$$

We introduce names $INF(U), G_{INF}$ with $x \varepsilon INF(U)$ meaning x to be an atomic object of the form $Inf(a)$ with $a \varepsilon U$, and $x \varepsilon G_{INF}$ meaning x to be a class of atomic objects, i.e., $x = Cls_{INF}(\Phi | INF(U))$ for some non–vacuous property Φ with $\Phi | INF(U)$ being its restriction to INF–atomic objects, and non–vacuous on those objects.

Rough inclusions we define below, are defined on the universe $INF(U)$. As a result of operations in this universe, some constructs are obtained that we interpret in the universe U.

To this end, we define the semantic operator [.], which to any object g over the universe $INF(U)$ (be it an atom or a class) assigns the class $[g] = Cls_U(x \varepsilon U : Inf(x) el_{INF} g)$.

It is obvious that $x el_U [INF(x)]$; we state a property satisfied by [.], viz.,

$$g el_{INF} h \Rightarrow [g] el_U [h], \tag{32}$$

that follows directly from definitions and the transitivity of el_{INF}.

3 Examples

We consider an information system (U, A) and we define two basic examples of rough inclusions on $INF(U)$, that would be convenient in the sequel.

We define rough inclusions incrementally, invoking ontology introduced above. Both inclusions defined below satisfy the assumption that el_{INF} coincides with $=_{INF}$ on atomic objects. Thus in both cases, mereological classes coincide with sets; this may simplify the analysis for the reader. However, our analysis in general does not require this assumption.

3.1 The Menger Rough Inclusion

Given (U, A), hence, the information universe $INF(U)$, for x, y with $x \varepsilon INF(U)$, $y \varepsilon INF(U)$, where $x =_{INF} Inf(u), y =_{INF} Inf(v)$ with $u, v \varepsilon U$, we define,

$$DIS_A(x, y) = \{a \in A : a(u) \neq a(v)\}, \tag{33}$$

and then we let $y\mu_r^M x$ if and only if

$$e^{(-\Sigma_{a \in DIS_A(x,y)} w_a)} \geq r, \tag{34}$$

where $w_a \in (0, \infty)$ is a weight associated with the attribute a. Then μ^M does satisfy (18)–(25) with $x, y, z \varepsilon INF(U)$. For $x, y, z \varepsilon INF(U)$, the rule,

$$\frac{y\mu_r x, x\mu_s z}{y\mu_{r \cdot s} z}, \tag{35}$$

holds.

To justify our claim (35), we may observe that $DIS(y, z) \subseteq DIS(y, x) \cup DIS(x, z)$, which implies by (34) that $y\mu_r^M x, x\mu_s^M z$ imply $y\mu_{r \cdot s}^M z$, i.e., (35) holds. It is a matter of straightforward checking that μ^M satisfies (18)–(25).

We now extend μ^M over pairs of the form x, g, where $x \varepsilon INF(U)$, $g \varepsilon G_{INF}$. We define μ^M in this case as follows:

$$x\mu_r^M g \Leftrightarrow \exists y.y \varepsilon U \wedge INF(y) el_{INF} g \wedge x\mu_r^M INF(y). \tag{36}$$

Then, the rough inclusion μ^M on pairs x, y or x, g, where $x, y \varepsilon INF(U)$, $g \varepsilon G_{INF}$, does satisfy (18)–(25).

Indeed, (18) and (25) are obviously true. For (23), $x\mu_1^M g$ implies that there exists $y \varepsilon INF(U), y el_{INF} g$ with $x\mu_1^M y$ hence $x el_{INF} y$ hence $x =_{INF} y$ so that $x el_{INF} g$. (24) follows on similar lines.

Finally, we define μ^M on pairs of the form g, h, with $g \varepsilon G_{INF}, h \varepsilon G_{INF}$, by means of the formula,

$$g\mu_r^M h \Leftrightarrow \forall x \varepsilon INF(U).[x el_{INF} g \Rightarrow \exists y \varepsilon INF(U).(y el_{INF} h \wedge x\mu_r^M y)]. \tag{37}$$

It is true that the extended by (37) most general form of μ^M does satisfy (18)–(25).

Indeed, (18) and (25) hold obviously; proof of (23) is as follows:
for $g, h \varepsilon G_{INF}$,

$$g\mu_1^M h, \tag{38}$$

if and only if

$$\forall x \varepsilon INF(U).[x el_{INF} g \Rightarrow \exists y \varepsilon INF(U).y el_{INF} h \wedge x\mu_1^M y], \tag{39}$$

if and only if

$$\forall x \varepsilon INF(U).[x el_{INF} g \Rightarrow \exists y \varepsilon INF(U).y el_{INF} h \wedge x =_{INF} y], \tag{40}$$

if and only if

$$\forall x\varepsilon INF(U).(xel_{INF}g \Rightarrow xel_{INF}h). \tag{41}$$

Now, from (41), the conclusion $gel_{INF}h$ follows by (21).

Proof for (24) goes on parallel lines: given $g\mu_1^M h$ and $k\mu_r^M g$, for any $x\varepsilon INF(U)$, where $xel_{INF}k$, there exist $y\varepsilon INF(U), yel_{INF}g$ with $x\mu_r^M y$; similarly, we find $z\varepsilon INF(U), zel_{INF}h$ with $y\mu_1^M z$. It follows by (35) that $x\mu_{r\cdot 1}^M z$ hence $x\mu_r^M z$ and thus, x being arbitrary, $k\mu_r^M h$.

A similar argument will show that the extended by (37) rough inclusion μ^M does satisfy the transitivity rule,

$$\frac{k\mu_r^M h, h\mu_s^M g}{k\mu_{r\cdot s}^M g}. \tag{42}$$

To prove (42), consider $k, h, g\varepsilon G_{INF}$; given $x\varepsilon INF(U), xel_{INF}k$, we find y, $z\varepsilon INF(U)$, $yel_{INF}h, zel_{INF}g$ with $x\mu_r^M y, y\mu_s^M z$, hence, by (35) we have $x\mu_{r\cdot s}^M z$ implying by arbitrariness of x that $k\mu_{r\cdot s}^M g$.

3.2 The Łukasiewicz Rough Inclusion

In notation of (33), we define a rough inclusion μ^L. Following the pattern of the discussion in (34), (35), the rough inclusion μ^L will be defined as follows.

For $x, y\varepsilon INF(U)$,

$$y\mu_r^L x \Leftrightarrow 1 - \frac{|DIS_A(x,y)|}{|A|} \geq r. \tag{43}$$

For $g, h\varepsilon\{INF(U), G_{INF}\}$:

$$(g\mu_r^L h) \Leftrightarrow \{[\forall x\varepsilon INF(U).(xel_{INF}g \Rightarrow \exists y\varepsilon INF(U).(yel_{INF}h \wedge x\mu_r^L y)]\}. \tag{44}$$

Following the lines of proofs of (18)–(25) in case of the Menger rough inclusion, and of (42) in which the rough inclusion μ^M is replaced with the rough inclusion μ^L, we obtain that *the Łukasiewicz rough inclusion* μ^L *does satisfy* (18)–(25) *with* $x, y, z\varepsilon\{INF(U), G_{INF}\}$, and the following rule,

$$\frac{k\mu_r^L g, g\mu_s^L h}{k\mu_{\otimes(r,s)}^L h}, \tag{45}$$

holds, where $\otimes(r, s)$ is the Łukasiewicz functor (10).

As with the Menger rough inclusion, it is sufficient to verify (18)–(25) with $x, y, z\varepsilon INF(U)$ which is straightforward and then to follow general lines outlined above.

The transitivity rule (45) deserves a proof. It is sufficient to assume $x\mu_r^L y$, $y\mu_s^L z, x\mu_t^L z$ with $x, y, z\varepsilon INF(U)$.

As $DIS(x, z) \subseteq DIS(x, y) \cup DIS(y, z)$, substituting into (43),

$$1 - t = \frac{|DIS(x,z)|}{|A|} \leq \frac{|DIS(x,y)|}{|A|} + \frac{|DIS(y,z)|}{|A|} = (1-r) + (1-s), \quad (46)$$

hence $t \geq r+s-1$ and as t is non–negative we have finally $t \geq max\{0, r+s-1\} = \otimes(r, s)$.

The proof in the general case goes on lines of the proof in case of Menger's rough inclusion.

We add a new rule about rough inclusions, which is satisfied by Menger as well as Łukasiewicz rough inclusions with $f = Prod$, $\otimes$ respectively.

The *f*-Transitivity Property

$$\frac{x\mu_r y, y\mu_s z}{x\mu_{f(r,s)} z} \quad (47)$$

holds with a t–norm f.

We say that μ is an *f–rough inclusion.*

Note, that it follows from general properties of t–norms, that for each t–norm, T, there exists a rough inclusion μ^T that satisfies the T–transitivity rule, i.e.,

$$\frac{k\mu_r^T g, g\mu_s^T h}{k\mu_{T(r,s)}^T h}, \quad (48)$$

see [22], Prop. 13.

Finally, we state for the record the obvious property of Menger as well as Łukasiewicz rough inclusions, viz. symmetry, of which no explicit use was made as of yet.

The Symmetry Property

$$x\mu_r y \Leftrightarrow y\mu_r x, \quad (49)$$

holds for $\mu \in \{\mu^M, \mu^L\}$.

3.3 Remarks on the Notion of an Element Induced by a Rough Inclusion

In cases of Menger as well as Łukasiewicz rough inclusions, formulas defining μ are employing cardinalities of sets $DIS_A(x, y)$; in consequence, the notion of an element el_{INF} induced by either of μ^M, μ^L coincides with the notion of being indiscernible: $xel_{INF}y$ if and only if $IND_A(u(x), u(y))$ if and only if $u(x) =_U u(y)$, where $x = INF(u(x)), y = INF(u(y))$.

At the cost of introducing an additional structure into value sets V_a, where $a \in A$, we may generate a richer el_{INF} relation. For instance, we may assume preference relations $<_a$ in sets V_a, i.e., orderings on value sets V_a, studied by some authors, e.g., Greco et al., see [6]. Then, we may let,

$$DIS_A^<(x, y) = \{a \in A : a(x) < a(y)\}, \quad (50)$$

and,

$$x\mu_r^{L,<}y \Leftrightarrow 1 - \frac{|DIS_A^{<}(x,y)|}{|A|} \geq r. \quad (51)$$

Alternatively, we may introduce the set

$$DIS_A^{\geq}(x,y) = \{a \in A : a(x) \geq a(y)\}, \quad (52)$$

hence,

$$DIS_A^{\geq}(x,y) = A \setminus DIS_A^{<}(x,y), \quad (53)$$

and, in consequence,

$$x\mu_r^{L,<}y \Leftrightarrow |DIS_A^{\geq}(x,y)| \geq r \cdot |A|. \quad (54)$$

Thus, the notion an element $el_{INF}^{<}$ is as follows: $xel_{INF}^{<}y$ if and only if $a(x) \geq a(y)$ for each a in A.

It is straightforward to verify that $\mu^{L,<}$ is a rough inclusion; only (23) may require a comment. In this case, we consider x, y, z with $x\mu_1^{L,<}y$ and $z\mu_r^{L,<}x$.

We have: $DIS_A^{\geq}(x,y) = A$, i.e., $a(x) \geq a(y)$ for every a in A, and

$$|DIS_A^{\geq}(z,x)| \geq r \cdot |A|.$$

As $a(z) \geq a(x)$ implies $a(z) \geq a(y)$, it follows that $DIS_A^{\geq}(z,x) \subseteq DIS_A^{\geq}(z,y)$, hence, $|DIS_A^{\geq}(z,y)| \geq |DIS_A^{\geq}(z,x)| \geq r \cdot |A|$, i.e., $z\mu_r^{L,<}y$.

In an analogous way, we may define the corresponding Menger rough inclusion $\mu_r^{M,<}$, i.e.,

$$x\mu_r^{M,<}y \Leftrightarrow e^{-\Sigma_{a \in DIS_A^{\geq}(x,y)} w_a} \geq r. \quad (55)$$

Let us observe that

$$e^{-\Sigma_{a \in DIS_A(x,y)} w_a)} = 1 - \frac{\Sigma_{a \in DIS_A(x,y)} w_a}{1!} + \frac{(\Sigma_{a \in DIS_A(x,y)} w_a)^2}{2!} - ..., \quad (56)$$

and, assuming that the average value of $|DIS_A(x,y)|$ is $\frac{|A|}{2}$, letting in (56) $w_a = \frac{1}{|A|}$, for each $a \in A$, we obtain that

$$e^{-\Sigma_{a \in DIS_A(x,y)} w_a)} = 1 - \frac{|DIS_A(x,y)|}{|A|}, \quad (57)$$

with an error of order 0.06. We may, therefore, base examples that follow on the Łukasiewicz rough inclusion as a fairly close estimate to the Menger rough inclusion.

We consider the information system $\mathcal{A}$ in Table 1.

For the information system $\mathcal{A}$, we may calculate values of the Łukasiewicz rough inclusion in Table 2; as μ^L is symmetric, we show only the upper triangle of values.

As already observed, the relation of *element* is equal to the identity $=_U$.

Table 1. The information system $\mathcal{A}$.

U	a_1	a_2	a_3	a_4
x_1	1	1	1	2
x_2	1	0	1	0
x_3	2	0	1	1
x_4	3	2	1	0
x_5	3	1	1	0
x_6	3	2	1	2
x_7	1	2	0	1
x_8	2	0	0	2

Table 2. μ^L for Table 2.

U	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	1	0.5	0.25	0.25	0.5	0.5	0.25	0.25
x_2	*	1	0.5	0.5	0.5	0.25	0.25	0.25
x_3	*	*	1	0.25	0.25	0.25	0.25	0.5
x_4	*	*	*	1	0.75	0.75	0.25	0.0
x_5	*	*	*	*	1	0.5	0	0
x_6	*	*	*	*	*	1	0.25	0.25
x_7	*	*	*	*	*	*	1	0.25
x_8	*	*	*	*	*	*	*	1

Table 3. $\mu^{<,L}$ for Table 2.

U	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	1	1	0.75	0.5	0.75	0.5	0.75	0.75
x_2	0.5	1	0.5	0.5	0.5	0.25	0.5	0.5
x_3	0.5	1	1	0.5	0.5	0.25	0.75	0.75
x_4	0.75	1	0.75	1	1	0.75	0.75	0.75
x_5	0.75	1	0.75	0.75	1	0.5	0.5	0.75
x_6	1	1	1	1	1	1	1	1
x_7	0.5	0.75	0.5	0.5	0.5	0.25	1	0.5
x_8	0.5	0.75	0.75	0.25	0.25	0.25	0.75	1

Taking the natural ordering of integers as the preference ordering in all sets V_a for $a = a_1, a_2, a_3, a_4$, we induce the rough inclusion $\mu^{L,<}$ in the universe of the system $\mathcal{A}$, whose values are shown in Table 3.

As we see, the rough inclusion $\mu^{L,<}$ is not symmetric, and we may expect a different structure of granules, as we pass to the subject of knowledge granulation.

3.4 Granulation of Knowledge

The class operator Cls may be recalled now. We make use of it in defining a notion of a granule in the rough mereological universe. Granular computing (GC) as a paradigm has been proposed by Lotfi Zadeh [28] and it has been studied intensively by rough and fuzzy set theorists, see [15], [12]. According to an informal, intuitive, definition, due to Zadeh, a granule is a collection (abstracting from the precise meaning of the term) of objects bound together by similarity, closeness, or some other constraints alike. Here, we construct granules on the basis of rough mereological similarity of objects. Due to regularity of assumptions about rough inclusions, we are able to obtain also some regular properties of granules.

4 Rough Mereological Granulation

We assume a rough inclusion μ on a mereological universe (U, el). For given $r < 1$ and $x \varepsilon U$, we let,

$$g_r(x) = Cls(\Psi_r), \tag{58}$$

where,

$$\Psi_r(y) \Leftrightarrow y\mu_r x. \tag{59}$$

The class $g_r(x)$ collects all atomic objects satisfying the class definition with the concept Ψ_r.

We will call the class $g_r(x)$ the *r–granule* about x; it may be interpreted as a neighborhood of x of radius r. We may also regard the formula $y\mu_r x$ as stating *similarity* of y to x (to degree r). We do not discuss here the problem of representation of granules; in general one may apply sets as the underlying representation structure. It will turn out, that in case of Menger as well as Łukasiewicz rough inclusions, granules may be represented as sets of objects satisfying Ψ_r.

From (18)–(25), the following general properties of granulation operators g_r may be deduced.

$$y\mu_r x \Rightarrow yelg_r(x), \tag{60}$$

$$x\mu_r y \wedge yelz \Rightarrow xelg_r(z), \tag{61}$$

$$\forall z.[zely \Rightarrow \exists w, q.(welz \wedge welq \wedge q\mu_r(x))] \Rightarrow yelg_r(x), \tag{62}$$

$$yelg_r(x) \wedge zely \Rightarrow zelg_r(x), \tag{63}$$

$$s \leq r \Rightarrow g_r(x)elg_s(x). \tag{64}$$

Property (60) follows by definition of g_r and class definition (19, 20). Property (61) is implied by (23) and (60). Property (62) follows by (21). Property (63) follows by transitivity of el. Property (64) is a consequence of (25).

Let us observe that in case of DIS_A – induced rough inclusions, $g_1(x) = x$ is the class of elements of x hence x itself. By (60–64), the system $\{g_r(x) : x \in U, r \in [0,1]\}$ is a neighborhood system for a weak topology on the universe U.

In the matter of example, we consider the information system of Table 1 along with values of rough inclusions $\mu^L, \mu^{L,<}$ given, respectively, in Tables 2, 3. Admitting $r = 0.5$, we list below granules of radii 0.5 about objects $x_1 - x_8$ in both cases. We denote with the symbol $g_i, g_i^<$, respectively, the granule $g_{0.5}(x_i)$ defined by $\mu_L, \mu_{L,<}$, respectively.

We have,

1. $g_1 = \{x_1, x_2, x_5, x_6\}$,
2. $g_2 = \{x_1, x_2, x_3, x_4, x_5\}$,
3. $g_3 = \{x_2, x_3, x_8\}$,
4. $g_4 = \{x_2, x_4, x_5, x_6\}$,
5. $g_5 = \{x_1, x_2, x_4, x_5, x_6\}$,
6. $g_6 = \{x_1, x_4, x_5, x_6\}$,
7. $g_7 = \{x_7\}$,
8. $g_8 = \{x_3, x_8\}$,

what provides an intricate relationship among granules: $_1, g_4, g_6 \subseteq g_5$, $g_8 \subseteq g_2$, g_2, g_5 incomparable by inclusion, g_7 isolated. We may contrast this picture with that for $\mu^{L,<}$.

1. $g_1^< = g_4^< = g_5^< = g_6^< = U$,
2. $g_2^< = g_3^< = g_7^< = U \setminus \{x_6\}$,
3. $g_8^< = \{x_1, x_2, x_3, x_7, x_8\}$,

providing a nested sequence of three distinct granules.

4.1 Granulation via the Menger Rough Inclusion

We now consider the Menger rough inclusion μ^M in which case the following additional properties of granules may be observed. The r-granule about x induced by μ^M will be denoted with the symbol g_r^M.

The following are additional properties of Menger granules,

$$x\varepsilon INF(U) \wedge xel_{INF}g_r^M(y) \Leftrightarrow x\mu_r^M y, \tag{65}$$

$$x\varepsilon INF(U) \wedge xel_{INF}g_r^M(y) \Rightarrow \forall t \in [0,1].g_t^M(x)el_{INF}g_{t\cdot r}^M(y). \tag{66}$$

For (65), assume that $xel_{INF}g_r^M(y)$ for $x\varepsilon INF(U)$; then, by class definition (19, 20), we find w, q such that $wel_{INF}x, wel_{INF}q, q\mu_r^M y$. Thus $w\mu_{1\cdot r}^M y$, hence $w\mu_r^M y$, and again, by $x\mu_1^M w$, it follows that $x\mu_r^M y$. Proof of (66) follows on the same lines.

4.2 Granulation via the Łukasiewicz Rough Inclusion

With the Łukasiewicz rough inclusion, analogous observations follow whose proofs are carried on the same lines. The following are additional properties of Łukasiewicz granules,

$$x\varepsilon INF(U) \wedge xel_{INF}g_r^L(y) \Leftrightarrow x\mu_r^L y, \tag{67}$$

$$x\varepsilon INF(U) \wedge xel_{INF}g_r^L(y) \Rightarrow \forall t \in [0,1].g_t^L(x)el_{INF}g_{\otimes(t,r)}^L(y). \tag{68}$$

We may sum up more suggestively (66, 68), viz., for $\mu \in \{\mu^M, \mu^L\}$ and $x\varepsilon INF(U)$, and sufficiently small positive ε,

$$xel_{INF}g_r(y) \Rightarrow g_{1-\varepsilon}(x)el_{INF}g_{r-\varepsilon}(y). \tag{69}$$

In the sequel, we discuss schemes for computing with granules of knowledge/computing with words leading to networks of units organized similarly to neural networks but performing their computations in different manner. Hybrid rough-neural computing schemes are recently given much attention, see [15]. Here, we propose a look at the schemes based purely on agents equipped with rough inclusions. Differentiability of the Menger type rough inclusions make them susceptible to learning by back-propagation, a subject we discuss in the final part of the following section.

5 Rough Mereological Granular Computing

We define an intelligent unit modelled on a classical perceptron [1] and then we develop the notion of a network of intelligent units of this type. The scheme presented here is a modification of the scheme proposed by Polkowski and Skowron in [22].

5.1 Rough Mereological Perceptron

We exhibit the structure of a *rough mereological perceptron* (RMP). It consists of an intelligent unit denoted ia. The input to ia is a finite set of connections $Link_{ia,in} = \{link_1, ..., link_m\}$; each $link_j$ has as the source the information universe $INF(U_j)$, of an information system $\mathcal{A}_j = (U_j, A_j)$, endowed with a rough inclusion μ_r^j. The output to ia is a connection $Link_{ia,out}$ to the information universe $INF(U_{ia})$, of an information system $\mathcal{A}_{ia} = (U_{ia}, A_{ia})$, equipped with the rough inclusion μ^{ia}.

The operation (function) realized in RMP is denoted with O_{ia}; thus, for every tuple $< x_1, ..., x_m >$, where $x_i \in INF(U_i)$, the object $x = O_{ia}(x_1, ..., x_m)$ $\in INF(U_{ia})$.

In each $INF(U_j)$ as well as in $INF(U_{ia})$, finite sets T_j, T_{ia} are selected, with the properties that,

$$t \in t_{ia} \Rightarrow \exists t_1 \in T_1, ..., t_m \in T_m.t = O_{ia}(t_1, .., t_m), \tag{70}$$

and,

$$\forall i, t_i \in T_i.\exists t \in T_{ia}.t = O_{ia}(t_1, ..., t_m). \tag{71}$$

When $t = O_{ia}(t_1, ..., t_m)$, we say that $\sigma =< t_1, ..., t_m, t >$ is an *admissible set of references*. The set of all admissible reference sets is denoted by Σ.

The operation of an RMP may be expressed in terms of the mapping ω_{ia} defined as follows.

For $\sigma =< t_1, ..., t_m, t >$, $r_1, ..., r_m \in [0, 1]$, $x = O_{ia}(x_1, ..., x_m)$, and $x_i \varepsilon INF(U_i)$ for $i = 1, 2, ..., m$, the mapping,

$$\omega_{ia} : \Sigma \times [0, 1]^m \rightarrow [0, 1], \tag{72}$$

is given by

$$\omega_{ia}(\sigma, r_1, .., r_m) = r = min\{s : x\mu_s^{ia} t\}, \tag{73}$$

whenever $x_i \mu_{r_i}^i t_i$ for $i = 1, 2, ..., m$.

5.2 Granular Computations

The mapping ω_{ia} may be factored through granule operator, i.e., (73) may be expressed as follows:

$$\omega_{ia}^g(\sigma, g_{r_1}(t_1), ..., g_{r_m}(t_m)) = g_{\omega_{ia}(\sigma, r_1, ..., r_m)}(t), \tag{74}$$

defining the factored mapping ω^g on granules.

Let us observe that the acting of RMP may as well be described as that of a *granular controller*, viz., the functor ω^g may be described via a decision algorithm consisting of rules of the form,

$$\mathbf{i}f\ g_{r_1}(t_1)\ \mathbf{a}nd\ g_{r_2}(t_2)\ \mathbf{a}nd\ \ldots\ \mathbf{a}nd\ g_{r_m}(t_m)\ \mathbf{t}hen\ g_r(t), \tag{75}$$

with $r = \omega_{ia}(\sigma, r_1, ..., r_m)$, where $\sigma =< t_1, ..., t_m, t >$.

It is worth noticing that the mapping ω_{ia} is defined from given information systems $\mathcal{A}_j, \mathcal{A}_{ia}$ and it is not subject to any arbitrary choice.

Composition of RMP's involves a composition of the corresponding mappings ω, viz., given $RMP_1, \ldots, RMP_k$, RMP with links to RMP being outputs from $RMP_1, \ldots, RMP_k$, each RMP_j having inputs $Link_j = \{link^j_1, .\,.., link^j_{k_j}\}$, $m = \Sigma^k_{j=1} k_j$, the composition,

$$IA = RMP \circ (RMP_1, RMP_2, ..., RMP_k), \tag{76}$$

of m inputs, does satisfy the formula

$$\omega_{IA} = \omega_{RMP} \circ (\omega_{RMP_1}, ..., \omega_{RMP_k}), \tag{77}$$

under the tacit condition that admissible sets of references are composed as well. Thus RMP's may be connected in networks subject to standard procedures, e.g., learning by back-propagation.

When learning by back-propagation (see [1], [20] for a detailed analysis), the Menger rough inclusion, or its modifications, e.g., the *gaussian rough inclusion*,

$$x \mu^G_r y \Leftrightarrow e^{-(\Sigma_{a \in DIS_A(x,y)} w_a)^2} \geq r, \tag{78}$$

are useful as they are defined by means of a differentiable function f, whose gradient is of the form,

$$\frac{\partial f}{\partial w} = f \cdot (-2 \cdot \sum w_a), \tag{79}$$

in case of μ^G.

5.3 Elementary Teams of Rough Mereological Perceptrons with the Menger Rough Inclusion

We consider a *rough mereological perceptron* (RMP), as defined in Sect. 5.1. It consists of an intelligent unit denoted ia whose input is a finite tuple $\overline{x} =< x_1, x_2, \ldots, x_k >$ of objects, see Sect. 5.1. We endow ia with the Menger rough inclusion $\mu^{M,ia}$ at ia.

The unit ia is also equipped with a set of *target objects* $T_{ia} \subset INF(U_{ia})$. Each object in T_{ia} represents a class of indiscernibility $IND_{A_{ia}}$ via the map INF.

The output of RMP is the granule of knowledge $res_{ia}(\overline{x}) = g_{r_{res}}(x)$ with the property that

$$r_{res} = max\{r : \exists y \in T_{ia}.O_{ia}(\overline{x})\mu^{M,ia}_r y\}, \tag{80}$$

i.e., we essentially classify $\overline{x}$ as the exact concept being the class of indiscernibility classes as close to the indiscernibility class of x as the closest target class, i.e., in degree r_{res}. Let us observe that this classification depends on the weight system $\{w_a : a \in A_{ia}\}$ chosen. We may also observe that we could also present the result of computation $res_{ia}(\overline{x})$ as the tuple $< g_{r_i}(t_i) : t_i \in T_{ia} >$ where $r_i = sup\{r : O_{ia}(\overline{x})\mu_r^{M,ia} t_i\}$ for each $i \leq |T_{ia}|$.

5.4 Rough Mereological Network

Now, we may consider a network of intelligent agents (NIA) organized in the manner of a feed-forward neural network, viz., we single out the following components in NIA.

1. The set *INPUT* of intelligent agents (*ia*'s) constituting the input layer of NIA. Each intelligent unit *ia* in *INPUT* acts as an RMP: it receives a tuple $\overline{x}_{ia}$ of objects from the stock of primitive objects (signals, etc.); it assembles $\overline{x}_{ia}$ into the object $O_{ia}(\overline{x}_{ia})$ in U_{ia}; it classifies $O_{ia}(\overline{x}_{ia})$ with respect to its target concepts in T_{ia} and outputs the result $res_{ia}(\overline{x}_{ia})$.
2. Sets $LEV_1, \ldots, LEV_k$ constituting consecutive inner layers of NIA with each *ia* in $\bigcup_i LEV_i$ acting as above.
3. The output to RMP denoted ia_{out} acting as each *ia* except that its computation result is the collective computation result of the whole network.

Although, in general, the structure of neural networks is modelled on directed acyclic graphs, we assume for simplicity, that NIA is ordered by a relation ρ into a tree. Thus ia_{out} is the root of the tree, $INPUT$ is the leaf set of the tree, and each LEV_i constitutes the corresponding level of pairwise non–communicating units in the tree.

Elementary Computations. To analyze computation mechanisms in NIA, we begin with *elementary computations* performed by each subtree of the form $nia = \{ia, ia_0, \ldots, ia_m\}$ with $ia_j \rho ia$ for each $j \leq m$, i.e., ia is the root unit in nia, and $ia_0, \ldots, ia_m$ are its daughter units.

We make one essential assumption about NIA, viz., we presume that after preliminary training, the target sets T_{ia}, T_{ia_j} have been coordinated, i.e.,

1. for each $t \in T_{ia}$, there exist $t_1 \in T_{ia_1}, \ldots, t_m \in T_{ia_m}$ such that $t = O_{ia}(t_1, \ldots, t_m)$;
2. each $t_j \in T_{ia_j}$ can be completed by $t_1 \in T_{ia_1}, \ldots, t_{j-1} \in T_{ia_{j-1}}, t_{j+1} \in T_{ia_{j+1}}, \ldots, t_m \in T_{ia_m}$ such that $t = O_{ia}(t_1, \ldots, t_m) \in T_{ia}$.

On the basis of these assumptions, we can describe elementary computations in terms of target sets. To this end, we introduce notions of *propagating functors* relative to sets of target concepts. The notion of a propagating functor was introduced in [22], and here we adopt it in a slightly changed form.

We will say that a set $\sigma = \{t, t_1, \ldots, t_m\}$ of target concepts, where $t \in T_{ia}, t_1 \in T_{ia_1}, \ldots, t_m \in T_{ia_m}$ is *admissible* when $t = O_{ia}(t_1, \ldots, t_m)$. For an admissible $\sigma = \{t, t_1, \ldots, t_m\}$ of target concepts, where $t \in T_{ia}, t_1 \in T_{ia_1}, \ldots, t_m \in T_{ia_m}$, the propagating functor $\Phi_{nia,\sigma}$ is defined as follows.

For $\overline{x} =< x_1 \in INF(U_{ia_1}), \ldots, x_m \in INF(U_{ia_m}) >$, we denote by $\overline{r}$ the tuple $< r_1, \ldots, r_m > \in \mathbf{R}^m$ defined by

$$\forall i.r_i = sup\{r : x_i \mu_r^{M,ia_i} t_i\}, \tag{81}$$

and we let

$$\Phi_{nia,\sigma}(\overline{r}) = sup\{s : O_{ia}(x_1, \ldots, x_m) \mu_s^{M,ia} t\}. \tag{82}$$

In terms of functors $\Phi_{nia,\sigma}$, reasoning and computations are carried out in NIA (see [22], [20] for details). As induced by the function f, $\Phi_{nia,\sigma}$ is piece–wise differentiable, allowing for back–propagation based learning in NIA (see [20]).

Computing with Words. The paradigm of computing with words [29] assumes that syntax of the computing mechanism is given as that of a subset of natural language, whereas the semantics is given in a formal computing mechanism involving numbers.

Let us comment briefly on how this idea may be implemented, exemplarily, in an RMP. Assume, there is given a set $\mathcal{N}$ of noun phrases $\{n_1, n_2, ..., n_m, n\}$ corresponding to information system universes $U_1, ..., U_m$. A set $\mathcal{ADJ}$ of adjective phrases is also given, and to each $\sigma \in \Sigma$, a set $adj_1, .., adj_m, adj$ is assigned.

Then the decision rule (75) may be expressed in the form,

$$\mathbf{i}f\ n_1\ \mathbf{is}\ adj_1\ \mathbf{a}nd\ \ldots \mathbf{a}nd\ n_m\ \mathbf{i}s\ adj_m\ \mathbf{t}hen\ n\ \mathbf{i}s\ adj. \tag{83}$$

The semantics of (83) is expressed in the form of (75). The reader will observe that (83) is similar in form to decision rules of a fuzzy controller, while the semantics is distinct.

Composition of RMP's as above is reflected in compositions of rules of the form (83) with semantics expressed by composed mappings ω.

In the light of analysis that begins in the next section, we may call the computation model presented above, a rough–fuzzy–neurocomputing model, see [15] for discussions of this topic.

6 Rough Set Approximations and Fuzzy Partitions and Equivalences Induced by Rough Inclusions

We assume that a mereological universe U is given with a rough inclusion μ of which we assume symmetry (49) and an f–transitivity property (47) with a t–norm f. We will construct in this setting rough respectively fuzzy universes derived from the rough inclusion in question, and represented by rough set approximations and by partitions and equivalence relations in the sense of fuzzy set theory.

In rough as well as fuzzy set–theoretic literature, many authors have devoted their attention to the problem of mutual relationships between the two theories and, in consequence, some notions of a rough –fuzzy set as well as a fuzzy–rough set have emerged. The reader will consult [2], [3], [5], [8], [9], [10], [25], as well as many chapters in this volume.

6.1 Rough Environment from Rough Inclusions

We first give an abstract description of our ideas in the rough case and then we explicate them in case of information systems and Menger/Łukasiewicz rough inclusions. We will here discern between universes $U, INF(U)$, which will bear on notation.

Assume a rough inclusion μ on a mereological universe $(INF(U), el_{INF})$ and a non–vacuous property M on classes over $INF(U)$; given a class x over $INF(U)$, we define its *lower* M, μ*–approximation* $\underline{x}^{\mu}_{M}$ as follows,

$$\underline{x}^{\mu}_{M} = Cls(g\varepsilon M : g\mu_1 x). \tag{84}$$

Then we have as a direct consequence of (21),

$$\underline{x^{\mu}}_{M} \ el \ x. \tag{85}$$

Indeed, for $y \ el_{INF}\underline{x^{\mu}}_{M}$ there exist w, q with

$$wel_{INF}y, wel_{INF}q, q\varepsilon M, q\mu_1 x, \tag{86}$$

hence, $qel_{INF}x$ by (18), and by the inference rule (21) it follows that,
$\underline{x^{\mu}}_{M} \ el_{INF}x$.

Again, applying the inference rule (21), we obtain,

$$\underline{x}^{\mu}_{M} \ =_{INF} \underline{\underline{x}^{\mu}_{M}}^{\mu}_{M}. \tag{87}$$

It is sufficient to verify that $\underline{x}^{\mu}_{M} \ el_{INF}\underline{\underline{x}^{\mu}_{M}}^{\mu}_{M}$, the converse being satisfied in virtue of (85). Given $yel_{INF}\underline{x}^{\mu}_{M}$, we have w, q with

$$wel_{INF}y, wel_{INF}q, q\varepsilon M, q\mu_1 x, \tag{88}$$

hence, $qel_{INF}\underline{x}^{\mu}_{M}$, i.e., $q\mu_1\underline{x}^{\mu}_{M}$ i.e. $qel_{INF}\underline{\underline{x}^{\mu}_{M}}^{\mu}_{M}$.

It follows by (21) that $\underline{x}^{\mu}_{M} \ el_{INF}\underline{\underline{x}^{\mu}_{M}}^{\mu}_{M}$.

We will say that x is M*–exact* in case there exists a non–empty $M_0 \subseteq M$ such that $x =_{INF} Cls(M_0)$.

Then, if

$$x = \underline{x}^{\mu}_{M}, \tag{89}$$

then x is M–exact.

Indeed, if $x = \underline{x}^{\mu}_{M}$ then $x = Cls(g\varepsilon M : g\mu_1 x)$ so it suffices to let $M_0(g) \Leftrightarrow g\varepsilon M \wedge y\mu_1 x$.

The converse also holds, i.e., if x is M–exact, $x = Cls(M_0)$ with some $M_0 \subseteq M$ then,

$$x =_{INF} \underline{x}^{\mu}_{M}. \tag{90}$$

Indeed, given $yel_{INF}x$, we have w, q with

$$wel_{INF}y, wel_{INF}q, q\varepsilon M_0, qel_{INF}x, \tag{91}$$

hence, $qel_{INF}\underline{x}^{\mu}_{M}$, and (21) implies that $xel_{INF}\underline{x}^{\mu}_{M}$, so (90) follows by (85).

Now, we define the *upper* M, μ*–approximation* $\overline{x}^{\mu}_{M}$ by letting

$$\overline{x}^{\mu}_{M} = Cls(g\varepsilon M : \exists yel_{INF}x.y\mu_1 g). \tag{92}$$

This case requires some assumptions about M. We will say that M is a *covering* in case the condition holds,

$$\forall x\varepsilon INF(U).\ \exists g\varepsilon M.\ x\mu_1 g. \tag{93}$$

We say that M is a *partition* in case the following condition holds,

$$\forall h, g\varepsilon M.(h \neq_{INF} g \Rightarrow ext(h, g)), \tag{94}$$

where

$$ext(h, g) \Leftrightarrow \neg[\exists z\varepsilon INF(U).\ (zel_{INF}h \wedge zel_{INF}g)]. \tag{95}$$

It is true that if M is a covering, then

$$xel_{INF}\overline{x}^{\mu}_{M}. \tag{96}$$

For $yel_{INF}x$, there exists $g\varepsilon M$ with $y\mu_1 g$ and then $yel_{INF}g, gel_{INF}\overline{x}^{\mu}_{M}$; by (21), $xel_{INF}\overline{x}^{\mu}_{M}$.

If M is a covering, then

$$\overline{x}^{\mu}_{M} =_{INF} \overline{\overline{x}^{\mu}_{M}}^{\mu}_{M}. \tag{97}$$

It suffices by (96) to verify the element inclusion of the right–hand side into the left–hand side. So, let $yel_{INF}\overline{\overline{x}^{\mu}_{M}}^{\mu}_{M}$. Thus there exist w, q with

$$wel_{INF}y, wel_{INF}q, qel_{INF}\overline{x}^{\mu}_{M}, \tag{98}$$

hence, there exist t, r, z with

$$tel_{INF}w, tel_{INF}r, r\varepsilon Mzel_{INF}x, zel_{INF}r, \tag{99}$$

and thus $rel_{INF}\overline{x}^{\mu}_{M}$, and as $tel_{INF}y$ by (98), (99), and transitivity of el_{INF}, the entities y, t, r do witness by (21) that $\overline{\overline{x}^{\mu}_{M}}^{\mu}_{M} el_{INF}\overline{x}^{\mu}_{M}$.

The question whether $x =_{INF} \overline{x}^{\mu}_{M}$ has a positive answer, too.

If M is a covering and a partition, and $x =_{INF} Cls(M_0)$ with some $\emptyset \neq M_0 \subseteq M$, then,

$$x =_{INF} \overline{x}^{\mu}_{M}. \tag{100}$$

It suffices to verify that $\overline{x}^{\mu}_{M} el_{INF} x$. To this end, we apply (21), and given $y el_{INF} \overline{x}^{\mu}_{M}$, we find w, q, z such that

$$w el_{INF} y, w el_{INF} q, q el_{INF} \overline{x}^{\mu}_{M}, q \varepsilon M, z el_{INF} x, z el_{INF} q. \tag{101}$$

As $x =_{INF} Cls(M_0)$, there exist u, r with

$$u el_{INF} z, u el_{INF} r, r \varepsilon M_0. \tag{102}$$

It follows that $u el_{INF} q, u el_{INF} r$ so $\neg ext(q, r)$ holds and thus by (95), $q =_{INF} r$. As $x =_{INF} Cls(M_0)$, $r el_{INF} x$ and from,

$$w el_{INF} r, w el_{INF} y, r el_{INF} x, \tag{103}$$

we infer by (21)) that $\overline{x}^{\mu}_{M} el_{INF} x$, concluding the proof.

In consequence of (89, 90, 100), when M is a covering and a partition, $\overline{x}^{\mu}_{M} = x = \underline{x}^{\mu}_{M}$ if and only if x is M–exact.

The above properties witness that rough mereological approximations share well–known [16] properties of rough set approximations.

We now should restore from rough mereological approximations $\underline{x}^{\mu}_{M}, \overline{x}^{\mu}_{M}$, rough set approximations $\underline{A}x, \overline{A}x$ over the universe U.

We first define the upper approximation. To this end, we make use of the semantic operator [.] subject to (32), and we let

$$\overline{x}^{M,\mu} = [\overline{INF(x)}^{\mu}_{M}]. \tag{104}$$

The following properties are true.

If M is a covering, then

$$x el_U \overline{x}^{M,\mu}. \tag{105}$$

If M is a covering, then

$$\overline{x}^{M,\mu} =_U \overline{\overline{x}^{M,\mu}}^{M,\mu}. \tag{106}$$

If M is a covering and a partition, then

$$x =_U \overline{x}^{M,\mu}, \tag{107}$$

whenever $x =_U [INF(x)]$ and $[INF(x)]$ is M–exact.

Property (105) follows from (32) and (96). Similarly, (106) follows from (32) and (97). For (107), if $INF(x)$ is M–exact, then

$INF(x) =_{INF} \overline{INF(x)}^{\mu}_{M}$ by (U3),

hence, $\overline{x}^{M,\mu} =_U [\overline{INF(x)}^{\mu}_{M}] =_U [INF(x)] =_U x$ by (32).

To define the lower approximation $\underline{x}_{M,\mu}$ to x, we make use of the upper approximation,

$$\underline{x}_{M,\mu} =_U (\overline{x^c}^{M,\mu})^c. \tag{108}$$

The following properties hold.

If M is a covering, then
$$\underline{x}_{M,\mu} el_U x. \tag{109}$$

If M is a covering, then
$$\underline{x}_{M,\mu} =_U \underline{\underline{x}_{M,\mu}}_{M,\mu}. \tag{110}$$

If $x =_U [INF(x)]$, $INF(x)]$ is M–exact, and M is a covering and a partition, then
$$x =_U \underline{x}_{M,\mu}. \tag{111}$$

Equation (109) holds as $x^c el_U \overline{x^c}^{M,\mu}$ by (105), hence
$$\underline{x}_{M,\mu} =_U (\overline{x^c}^{M,\mu})^c el_U x^{cc} =_U x.$$

Equation (110) results from (106), as
$$\underline{\underline{x}_{M,\mu}}_{M,\mu} =_U (\overline{(\underline{x}_{M,\mu})^c}^{M,\mu})^c =_U (\overline{(\overline{x^c}^{M,\mu})^{cc}}^{M,\mu})^c =_U$$
$$(\overline{(\overline{x^c}^{M,\mu})}^{M,\mu})^c =_U (\overline{x^c}^{M,\mu})^c =_U \underline{x}_{M,\mu}.$$

Equation (111) follows from (107), as $x^c =_U [INF(x^c)]$ and $INF(x^c)$ is M–exact hence $x^c =_U \overline{x^c}^{M,\mu}$ and thus $x =_U x^{cc} =_U (\overline{x^c}^{M,\mu})^c = \underline{x}_{M,\mu}$.

6.2 The Case of Menger, Łukasiewicz Rough Inclusions

We insert here a short interlude on the approximation theme using the Menger or the Łukasiewicz rough inclusions to define granules denoted in this case with the symbol $g_r(x)$. As the collection,
$$M = \{g_{1-\varepsilon}(x) : x\varepsilon U\}, \tag{112}$$

with ε sufficiently small is a covering and a partition in $INF(U)$, and
$$g_{1-\varepsilon}(x) =_{INF} Inf(x), \tag{113}$$

it follows that $\overline{x}^{\mu}_M = x = \underline{x}^{\mu}_M$. Applying the operator [.], we reach the conclusion that
$$\underline{x}_{M,\mu} = \underline{A}x. \tag{114}$$

Similarly, for the upper approximation,
$$\overline{x}^{M,\mu} = \overline{A}x. \tag{115}$$

For the argument, it suffices to notice that $[Inf(x)] = [x]_{IND(A)}$.

6.3 Fuzzy Partitions and Equivalences from Rough Inclusions

Because $y\mu_1 x$ is equivalent to $y\ el\ x$, we may interpret $y\mu_r x$ as a statement of fuzzy membership; we will write $\mu_x(y) = r$ to stress this interpretation.

Clearly, by its definition, a rough inclusion is a relation, or, it may be regarded as a generalized fuzzy membership function, that takes intervals of the form $[0, w]$ as its values, leading to a higher–level fuzzy set. That follows from (25). Hence, inequalities like $a \geq \tau(x, y)$ are interpreted as follows: the value of a belongs in the interval - the value of $\tau(x, y)$. We should also bear in mind that rough inclusions reflect information content of the underlying information system.

Thus, we may say, that rough inclusions induce globally a family of fuzzy sets $\{x : xel_{INF}INF(U)\}$ with fuzzy membership functions $\{\mu_x : xel_{INF}INF(U)\}$, in the above sense.

We assume additionally for the considered rough inclusion μ the f-transitivity property (47) with a t–norm f. Let us consider a relation τ on U defined as follows,

$$x\tau_r y \Leftrightarrow x\mu_r y \wedge y\mu_r x, \tag{116}$$

hence, τ_r is for each r a *tolerance relation*. The following properties hold,

$$x\tau_1 x, \tag{117}$$

$$x\tau_r y \Leftrightarrow y\tau_r x, \tag{118}$$

$$x\tau_r y \wedge y\tau_s z \Rightarrow x\tau_{f(r,s)} z. \tag{119}$$

We will write $\tau(x, y)$ instead of $\chi_{x,\tau}(y)$ in cases when we treat τ as a fuzzy set, except for cases when the latter notation is necessary.

We may paraphrase (117)–(119) in terms of the new notation,

$$\tau(x, x) = 1, \tag{120}$$

$$\tau(x, y) = \tau(y, x), \tag{121}$$

$$\tau(x, z) \geq f(\tau(x, y), \tau(y, z)). \tag{122}$$

Thus, τ is by (120), (121), and (122), an f-*fuzzy similarity*, see [27].

Following [27], we may define *similarity classes* $[x]_\tau$ as fuzzy sets satisfying,

$$\chi_{[x]_\tau}(y) = \tau(x, y). \tag{123}$$

The following are true,

$$\chi_{[x]_\tau}(x) = 1, \tag{124}$$

$$f(\chi_{[x]_\tau}(y), \tau(y, z)) \leq \chi_{[x]_\tau}(z), \tag{125}$$

$$f(\chi_{[x]_\tau}(y), \chi_{[x]_\tau}(z)) \leq \tau(y, z). \tag{126}$$

Equations (124, 125, 126) follow by corresponding properties of τ given by (117), (118), and (119).

Finally, the family $\{[x]_\tau : x \subseteq U\}$ does satisfy the requirements to be a *fuzzy partition* [5], [7], [23], [27], viz.,

$$\forall x \exists y . \chi_{[x]_\tau}(y) = 1, \tag{127}$$

$$[x]_\tau \neq [z]_\tau \Rightarrow \; max_y\{min\{\chi_{[x]_\tau}(y), \chi_{[z]_\tau}(y)\}\} < 1, \tag{128}$$

$$\bigcup_x [x]_\tau \times_f [x]_\tau = \tau, \tag{129}$$

where $A \times_f B$ denotes the fuzzy set defined via

$$\chi_{A \times_f B}(u, v) = f(\chi_A(u), \chi_B(v)) \tag{130}$$

and $\bigcup_x$ denotes the supremum over all values of x.

Indeed, (127), (128), (129) follow directly from properties of τ. For instance, (128) is justified as follows: if there was y with $\tau(x, y) = 1 = \tau(z, y)$, we would have $\tau(x, z), \tau(z, x) \geq f(1, 1) = 1$ hence $x =_U z$.

For (129), on one hand, given x, y, z, we have,

$$f(\tau(x, y), \tau(x, z)) = f(\tau(y, x), \tau(x, z)) \leq \tau(y, z)$$

by (122), hence, $\bigcup_x [x]_\tau \times_f [x]_\tau(y, z) \leq \tau(y, z)$.

On the other hand, letting $x = y$, we have,

$$[x]_\tau \times_f [x]_\tau(y, z) = [y]_\tau \times_f [y]_\tau(y, z) = f(\tau(y, y), \tau(y, z)) = f(1, \tau(y, z) = \tau(y, z)$$

by (120) and the property of every t–norm T that $T(1, u) = u$. Hence, $\bigcup_x [x]_\tau \times_f [x]_\tau(y, z) \leq \tau(y, z)$, which gives (129).

7 Conclusion

We may conclude our discussion; we have shown that rough mereological inclusions in the universe of an information system induce rough set approximations as well as fuzzy equivalence relations and partitions thus creating both rough as well as fuzzy framework. This concerns in particular the Menger as well as the Łukasiewicz rough inclusions, our exemplary rough inclusions.

References

1. C. Bishop. *Neural Networks for Pattern Recognition.* Clarendon, Oxford, repr. 1997.
2. G. Cattaneo. Generalized rough sets. Preclusivity fuzzy–intuitionistic (BZ) lattices. *Studia Logica*, **58**, 1997, pp. 47–77.
3. D. Ciucci. *Algebraic Structures for Rough Sets and Many–Valued Logics.* Manuscript. Univ. di Milano-Bicocca, 2003.
4. B. L. Clarke. A calculus of individuals based on "Connection". *Notre Dame J. Formal Logic*, **22**, 1981, pp. 204–218.

5. D. Dubois and H. Prade. Putting rough sets and fuzzy sets together, in: R. Słowiński Ed. *Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory.* Kluwer, 1992, pp. 203–232.
6. S. Greco, B. Matarazzo, R. Słowiński, and J. Stefanowski. An algorithm for induction of decision rules consistent with the dominance principle. *Lecture Notes in Artificial Intelligence* **2005**, Springer Verlag, Berlin, 2001, pp. 304–313.
7. U. Hoehle. Quotients with respect to similarity relations. *Fuzzy Sets and Systems* **27**, 1988, pp. 31–44.
8. R. Intan, Y. Y. Yao and M. Mukaidono. Generalization of rough sets using weak fuzzy similarity relations. In:[9], pp. 37–46.
9. M. Inuiguchi, S. Hirano and S. Tsumoto Eds. *Rough Set Theory and Granular Computing*, Physica, Heidelberg, 2003.
10. M. Inuiguchi and T. Tanino. Two directions toward generalization of rough sets. In: [9], pp. 47–58.
11. S. Leśniewski. On the foundations of mathematics. *Topoi* **2**, 1982, pp. 7–52.
12. T.Y. Lin. Granular computing: Fuzzy logic and rough sets. In: [29], pp. 183–200.
13. C.-H. Ling. Representation of associative functions. *Publ. Math. Debrecen* **12**, 1965, pp. 189–212.
14. E. Marczewski. A general scheme of independence in mathematics. *Bull. Polish Acad. Math. Sci.* **6**, 1958, pp. 731–736.
15. S. K. Pal, L. Polkowski, and A. Skowron Eds. *Rough–Neurocomputing. Techniques for Computing with Words.* Springer Verlag, in print.
16. Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning about Data.* Kluwer, 1992.
17. Z. Pawlak. Rough sets. *Intern. J. Comp. Inf. Sci.* **11**, 1982, pp. 341–356.
18. M, Novotný and Z. Pawlak. On rough equalities. *Bull. Polish Acad. Sci. Math.* **33**, 1985, pp. 99–104.
19. Z. Pawlak and A. Skowron. Rough membership functions, in: R. R. Yager, M. Fedrizzi, and J. Kacprzyk Eds. *Advances in the Dempster-Schafer Theory of Evidence.* Wiley, 1994, pp. 251–271.
20. L. Polkowski. A rough–neural computation model based on rough mereology. In: [15], pp. 85–108.
21. L. Polkowski. *Rough Sets. Mathematical Foundations.* Physica, Heidelberg, 2002.
22. L. Polkowski, A. Skowron. Rough mereology: a new paradigm for approximate reasoning. *International Journal of Approximate Reasoning* **15: 4**, 1997, pp. 333–365.
23. L. Valverde. On the structure of F–indistinguishability operators. *Fuzzy Sets and Systems* **17**, 1985, pp. 313–328.
24. A. N. Whitehead. *An Enquiry Concerning the Principles of Natural Knowledge.* Cambridge, 1919.
25. Y. Y. Yao. Generalized rough set models. In: L. Polkowski and A. Skowron Eds. *Rough Sets in Knowledge Discovery 1*, Physica, Heidelberg, 1998, pp. 286–320.
26. L. A. Zadeh. Fuzzy sets. *Information and Control* bf 8, 1965, pp. 338–353.
27. L. A. Zadeh. Similarity relations and fuzzy orderings. *Information Sciences* **3**, 1971, pp. 177–200.
28. L. A. Zadeh. Fuzy sets and information granularity. In: M. Gupta, R. Ragade, and R. Yager Eds. *Advances in Fuzzy Set Theory and Applications*, North Holland, Amsterdam, 1979, pp. 3–18.
29. L. A. Zadeh and J. Kacprzyk Eds. *Computing with Words in Information/Intelligent Systems* **1**, Physica, Heidelberg, 1999.