Multi-Adjoint Fuzzy Rough Sets

Chris Cornelis[a,1], Jesús Medina[b,2], Nele Verbiest[a]

[a] *Department of Applied Mathematics and Computer Science, Ghent University, Gent, Belgium*
[b] *Department of Mathematics, University of Cádiz, Spain*

Abstract

In this paper, we present an extension of the well-known implicator/t-norm based fuzzy rough set model based on a family of adjoint pairs. Rather than using a fixed implicator/t-norm pair to compute lower and upper approximations, our model allows to differentiate these fuzzy logical connectives according to the objects used in the calculation. We verify mathematical properties of the model, relate it to previous work in multi-adjoint property-oriented concept lattices, and discuss its possible use in data reduction.

1. Introduction

Fuzzy sets and rough sets model two complementary characteristics of imperfect knowledge: while the former allow that objects belong to a set or relationship to a given degree, the latter provide approximations of concepts when the available information is incomplete. A first definition of fuzzy rough sets was given in the eighties [5], and since then numerous hybrid models have been proposed, see e.g. [9, 10, 11, 12]. The most commonly used approach is the implicator/t-norm based model: given a fuzzy set $A\colon X \to [0,1]$ and a fuzzy relation R in X, $R\colon X \times X \to [0,1]$, the lower and upper approximation, $\downarrow, \uparrow\colon [0,1]^X \to [0,1]^X$, of A under R are defined by

$$(R{\downarrow}A)(y) = \inf_{x \in X} \mathcal{I}(R(x,y), A(x)) \tag{1}$$

$$(R{\uparrow}A)(y) = \sup_{x \in X} \mathcal{T}(R(x,y), A(x)) \tag{2}$$

where $\mathcal{T}$ is a t-norm and $\mathcal{I}$ a fuzzy implication. In this paper, we also assume that R is a fuzzy tolerance (i.e, reflexive and symmetric) relation and that $(\mathcal{I}, \mathcal{T})$ satisfies the adjoint property, that is, for all x, y and z in $[0,1]$

$$\mathcal{T}(x,y) \le z \quad \text{if and only if} \quad x \le \mathcal{I}(y,z) \tag{3}$$

Email addresses: `Chris.Cornelis@UGent.be` (Chris Cornelis), `jesus.medina@uca.es` (Jesús Medina), `Nele.Verbiest@UGent.be` (Nele Verbiest)

[1] Supported by Research Foundation–Flanders.

[2] Partially supported by Junta de Andalucía grant P09-FQM-5233, and by the EU (FEDER), and the Spanish Science and Education Ministry (MEC) under grant TIN2009-14562-C05-03.

Hence, $\mathcal{I}$ is the residuated implication of $\mathcal{T}$, and $(\mathcal{I}, \mathcal{T})$ is an adjoint pair [2].

The approximations given by (1) and (2) are static in a sense that the same adjoint pair is used regardless of x and y. However, it may be interesting to differentiate these operators depending on the objects under consideration. In this paper, we propose such a model in Section 2, and investigate which properties of classical rough sets are preserved. In Section 3, we investigate its relationship to previous work on multi-adjoint property-oriented concept lattices [7] and in Section 4 we motivate the practical use of the model in a data reduction task where the adjoint pair is determined by the relative noise level of the objects under consideration.

2. Model and Properties

Assume that for each x and y in X, there exists an adjoint pair $(\mathcal{I}^{x,y}, \mathcal{T}^{x,y})$, where $\mathcal{T}^{x,y}$ is a t-norm and $\mathcal{I}^{x,y}$ its residuated implication. The order matters, such that $(\mathcal{I}^{x,y}, \mathcal{T}^{x,y}) = (\mathcal{I}^{y,x}, \mathcal{T}^{y,x})$ does not necessarily hold. We can define the lower and upper approximation of A under R as

$$(R{\downarrow}A)(y) = \inf_{x \in X} \mathcal{I}^{x,y}(R(x,y), A(x)) \quad (4)$$

$$(R{\uparrow}A)(y) = \sup_{x \in X} \mathcal{T}^{x,y}(R(x,y), A(x)) \quad (5)$$

Clearly, if all adjoint pairs are equal, (1) and (2) are recovered. Table 1 lists elementary properties of the model w.r.t. monotonicity, interaction with set operations, and iterative application. It is interesting that the first five properties hold under the same conditions as (1) and (2). Property 6 is in general not necessarily fulfilled, even when R satisfies $\min$-transitivity.

Table 1: Properties of multi-adjoint approximations (R is a fuzzy tolerance relation).

	Property	Conditions
1.	$R{\uparrow}A = co_{\mathcal{N}}(R{\downarrow}(co_{\mathcal{N}}A))$ $R{\downarrow}A = co_{\mathcal{N}}(R{\uparrow}(co_{\mathcal{N}}A))$	$\mathcal{N}$ involutive
2.	$R{\downarrow}A \subseteq A \subseteq R{\uparrow}A$	
3.	$A \subseteq B \Rightarrow \begin{cases} R{\downarrow}A \subseteq R{\downarrow}B \\ R{\uparrow}A \subseteq R{\uparrow}B \end{cases}$	
4.	$R{\downarrow}(A \cap B) = R{\downarrow}A \cap R{\downarrow}B$ $R{\uparrow}(A \cap B) \subseteq R{\uparrow}A \cap R{\uparrow}B$	$(A{\cap}B)(x) = \min(A(x), B(x))$
5.	$R{\downarrow}(A \cup B) \supseteq R{\downarrow}A \cup R{\downarrow}B$ $R{\uparrow}(A \cup B) = R{\uparrow}A \cup R{\uparrow}B$	$(A{\cup}B)(x) = \max(A(x), B(x))$
6.	$R{\uparrow}(R{\downarrow}A) = R{\downarrow}(R{\downarrow}A) = R{\downarrow}A$ $R{\uparrow}(R{\uparrow}A) = R{\downarrow}(R{\uparrow}A) = R{\uparrow}A$	does not hold

3. Related Work

Property-oriented concept lattices [4] arise as an extension of rough set theory where two different sets, namely the set of objects and the set of attributes, are considered.

Recently this theory has been generalized to a fuzzy environment following the philosophy of the mult-adjoint paradigm [7, 8]. The authors worked in a general non-commutative environment and this naturally led to the consideration of several adjoint triples as the main building blocks of a multi-adjoint concept lattice.

Adjoint triples are more general than the adjoint pairs considered in this paper where, for example, the residuated conjunctor of an adjoint triple does not need to be commutative and it is defined on two complete lattices $(L_1, \preceq_1)$, $(L_2, \preceq_2)$, and a poset $(P, \leq)$. For more details see [6, 7].

In order to define the extension of the lower and upper approximations introduced in rough sets, a *context* must be fixed. A context is a tuple (A, B, R, σ) such that A and B are non-empty sets (usually interpreted as attributes and objects, respectively), R is a P-fuzzy relation $R\colon A \times B \to P$ and $\sigma\colon B \to \{1, \dots, n\}$ is a mapping which associates any element in B with some particular adjoint triple in the frame. Note that, in the expressions (4) and (5), an adjoint pair is associated with a pair of elements instead of only one.

Now, given $g \in L_2^B$, and $f \in L_1^A$, we define the following mappings: ${}^{\uparrow_\pi}\colon L_2^B \to L_1^A$, ${}^{\downarrow^N}\colon L_1^A \to L_2^B$:

$$g^{\uparrow_\pi}(a) = \sup\{R(a, b) \mathbin{\&_b} g(b) \mid b \in B\} \tag{6}$$

$$f^{\downarrow^N}(b) = \inf\{f(a) \nwarrow_b R(a, b) \mid a \in A\} \tag{7}$$

Clearly, considering $A = X$ and $[0, 1]$ as the complete lattices, these definitions are generalizations of the lower and upper approximations (4) and (5), if the adjoint pairs are associated with only one object.

Therefore, although the multi-adjoint fuzzy rough sets may be a particular case of a multi-adjoint property-oriented concept lattice, in this new environment pairs of objects are associated with adjoint pairs which will be an important characteristic in the application to feature selection. Moreover, more relevant properties can be proved.

4. Application to Feature Selection

Fuzzy rough sets have been applied to feature selection extensively, see e.g. [1]. We consider a decision system $(X, \mathcal{A} \cup \{d\})$ that consists of a universe of objects $X = \{x_1, \dots, x_m\}$ and a set of attributes $\mathcal{A} = \{a_1, \dots, a_n\}$ together with a fixed decision attribute d. Any x in X has a value $a(x)$ for each attribute a in $\mathcal{A}$ that can be real of nominal. We assume that $d(x)$, the class label of x, is always nominal. We also assume that a fuzzy tolerance relation R_B in X is defined for every $B \subseteq \mathcal{A}$. The B-positive region is then given as, for y in X,

$$POS_B(y) = (R_B{\downarrow}[y]_d)(y) = \inf_{x \in X} \mathcal{I}(R_B(x, y), [y]_d(x)) \tag{8}$$

where $[y]_d(x)$ is 1 if x has the same class label as y and 0 otherwise. A set $B \subseteq \mathcal{A}$ is called a fuzzy decision reduct if $|POS_B| = |POS_{\mathcal{A}}|$ [3], and there are no proper subsets of B with this

[3] The cardinality of a fuzzy set C in X is calculated by $|C| = \sum_{x \in X} C(x)$.

property. Given a fuzzy decision reduct B, the decision system can be reduced such that all attributes outside B are eliminated.

This feature selection process is very effective, but is also subject to the impact of noise. Since the lower approximation definition contains an infimum, a change in a single object can drastically affect the positive region. Many researchers have tried to tackle this problem by replacing the infimum by other operations to reduce the impact of "noisy" objects, including the VQRS model, OWA-based fuzzy rough sets, soft fuzzy rough sets, etc. In what follows, we argue that, given an $X \to [0,1]$ function α which assigns to each object its relative noise level (0 meaning no noise, 1 totally noisy), the multi-adjoint fuzzy rough set model can be invoked.

Given a family $(\mathcal{I}^\alpha)_{\alpha\in[0,1]}$ of fuzzy implications increasing in α, we may redefine (8) by

$$POS_B(y) = (R_B{\downarrow}[y]_d)(y) = \inf_{x\in X} \mathcal{I}^{\max(\alpha(x),\alpha(y))}(R_B(x,y),[y]_d(x)) \tag{9}$$

In this framework, if x is noisy (high $\alpha(x)$), then it will lead to a larger fuzzy implication value in (9), hence its impact on the computation of $POS_B(y)$ will be less since the infimum is taken over all objects in X. On the other hand, if y is noisy (high $\alpha(y)$), its membership to the positive region will be increased. In this way, a noisy object can be more easily distinguished from outliers in the data which are assumed to be correct; in the classical approach of (1), both would have a low positive region membership.

In our current experimentation on benchmark datasets, we are working with the family $(\mathcal{I}^\alpha)_{\alpha\in[0,1]}$ of Łukasiewicz implications defined by $\mathcal{I}^\alpha(x,y) = \sqrt[1+\alpha]{\min(1, 1-x^{1+\alpha}+y^{1+\alpha})}$, and noise levels determined by means of rough membership functions. Other setups will be tried in the future.

References

[1] C. Cornelis, R. Jensen, G. Hurtado, D. Slezak, Attribute Selection with Fuzzy Decision Reducts, Information Sciences **180(2)**, p. 209–224, 2010.

[2] R. P. Dilworth, M. Ward, Residuated lattices, Transactions of the American Mathematical Society **45**, p. 335–354, 1939.

[3] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems **17**, p. 91–209, 1990.

[4] I. Düntsch, G. Gediga, Approximation Operators in Qualitative Data Analysis, Theory and Applications of Relational Structures as Knowledge Instruments, Lect. Notes in Computer Science, **2929**, p. 214-230, 2003.

[5] L. Fariñas del Cerro, H. Prade, Rough sets, twofold fuzzy sets and modal logic—Fuzziness in indiscernibility and partial information, The Mathematics of Fuzzy Systems (A. Di Nola, A.G.S. Ventre, eds.), Verlag TUV Rheinland, Köln, p. 103–120, 1986.

[6] J. Medina, Towards multi-adjoint property-oriented concept lattices, Lect. Notes in Artificial Intelligence **6401**, p. 159–166, 2010.

[7] J. Medina, M. Ojeda-Aciego, J. Ruiz-Calviño, Formal concept analysis via multi-adjoint concept lattices, Fuzzy Sets and Systems **160(2)**, p. 130–144, 2009.

[8] J. Medina, M. Ojeda-Aciego, P. Vojtáš, Multi-adjoint logic programming with continuous semantics, Lect. Notes in Artificial Intelligence **2173**, p. 351–364, 2001.

[9] N.N. Morsi, M.M. Yakout, Axiomatics for fuzzy rough sets, Fuzzy sets and Systems **100**, p. 327–342, 1998.

[10] A. Nakamura, Fuzzy rough sets, Note on Multiple-Valued Logic in Japan **9**, p. 1–8, 1988.

[11] S. Nanda, S. Majumdar, Fuzzy rough sets, Fuzzy Sets and Systems **45**, p. 157–160, 1992.

[12] A.M. Radzikowska, E.E. Kerre, A comparative study of fuzzy rough sets, Fuzzy Sets and Systems **126**, p. 137–156, 2002.