Fundamenta Informaticae XXI (2001) 1001–1025
IOS Press

Rough Sets, Coverings and Incomplete Information

Inés Couso
Universidad de Oviedo, Spain
e-mail: couso@uniovi.es

Didier Dubois
IRIT, CNRS & Université de Toulouse, France
e-mail: dubois@irit.fr

Abstract. Rough sets are often induced by descriptions of objects based on the precise observations of an insufficient number of attributes. In this paper, we study generalizations of rough sets to incomplete information systems, involving imprecise observations of attributes. The precise role of covering-based approximations of sets that extend the standard rough sets in the presence of incomplete information about attribute values is described. In this setting, a covering encodes a set of possible partitions of the set of objects. A natural semantics of two possible generalisations of rough sets to the case of a covering (or a non transitive tolerance relation) is laid bare. It is shown that uncertainty due to granularity of the description of sets by attributes and uncertainty due to incomplete information are superposed, whereby upper and lower approximations themselves (in Pawlak's sense) become ill-known, each being bracketed by two nested sets. The notion of measure of accuracy is extended to the incomplete information setting, and the generalization of this construct to fuzzy attribute mappings is outlined.

Keywords: Rough sets, Possibility theory, Covering, fuzzy sets.

Address for correspondence: IRIT, CNRS & Université de Toulouse, France
e-mail: dubois@irit.fr

1. Introduction

Rough sets [23] are often induced by descriptions of objects based on the precise observations of attributes. The limited expressivity of the language defined by these attributes prevents sets of objects defined in extension from being precisely described in this language. Only upper and lower approximations of sets can be characterised in terms of unions of equivalence classes of the equivalence relation defined by the attributes. On the other hand, sets of objects defined by a property expressed in the language of the attributes may be ill-known when the attribute values of objects are imprecise. In this case, the attributes become set-valued mappings. However, these sets represent imprecision, i.e. they contain mutually exclusive values, one of which is the actual attribute value of the object under concern. Again in this case, a set of objects is only approached via an upper and a lower approximation, that is the sets of objects that possibly and necessarily satisfy the property, respectively [8].

This paper describes the situation where both sources of uncertainty are present. Orlowska and Pawlak [20] were among the first to consider the possibility of imprecision in information systems. These authors interpreted overlapping sets containing the attribute values of objects as a form of similarity between them. Two objects are related if they are potentially indistinguishable, i.e. their attribute values are possibly the same. They noticed that from ill-known attribute values, this relation on the set of objects is reflexive and symmetric and not transitive. It is no longer an equivalence relation.

Most of the research devoted to information systems in the eighties and the nineties focused on approximations of sets of objects based on complete information systems inducing equivalence classes of indiscernible objects. In this case the only source of uncertainty is the lack of a sufficient number of attributes. An early exception is Grzymala-Busse [13] who proposed to replace each item with missing attribute values in a data table by precise items bearing all possible values of the attribute, thus turning incompleteness into inconsistency. Research on the hybrid situation, putting together indiscernibility of objects and incomplete knowledge about attribute values, was revived in the late nineties (see [19] for a survey). For instance Kryszkiewicz [15] considered a special case where attribute values are either precisely known or totally unknown. He generalized the definition of rough sets turning equivalence classes of objects into sets of similar objects forming a covering.

The idea of defining abstract extensions of rough sets, replacing the partition associated with an equivalence relation by a covering has attracted many authors, the first of whom seems to be Pomykala [24]. The latter proposed two natural definitions of upper and lower approximations of a set of objects: one can extend the lower approximation of a set as the union of covering elements contained in the set, and the upper approximation as its dual via complementation; or extend the upper approximation of a set as the union of covering elements intersecting the set, and the lower approximation as its dual via complementation.These definitions were also discussed by Yao [31]. Many other definitions of covering-based rough sets have been proposed (see the extensive recent review by Samanta and Chakraborty [26]).

In this paper we try to bridge the gap between generalized rough sets based on coverings, and the possible-world approach to incomplete information. Each set in a covering is then the upper inverse image [7] of an attribute value through a multiple-valued mapping describing incomplete knowledge on attribute values. It is the set of objects that are possibly indistinguishable. Due to the presence of two sources of uncertainty, the rough set upper and lower approximations of a set are ill-known, each being bracketed by two nested sets. This interpretive setting leads us to choose, among possible covering-based rough sets definitions, some of them that look more appropriate than others.

The paper is organized as follows. Section 2 provides the formal setting and notations. Section

3 explains how incomplete information on attributes generates a covering of the set of objects and a tolerance relation and recalls some generalizations of rough sets based on coverings. Section 4 shows how upper and lower approximations of subsets of objects naturally emerge from the covering thus constructed. Moreover, by adopting a possible-world approach, viewing multiple-valued mappings as sets of selection functions, coverings are shown to be interpreted as imprecisely-known partitions. It enables some covering-based extensions of rough sets to be properly interpreted.

In Section 5, we also study what rough probability [22] becomes in such a setting. We obtain imprecise evaluations of Pawlak's upper and lower quality functions [21] and of the accuracy index of an approximation of a set. We also consider uncertainty measures, actually belief and plausibility functions, induced on the set of objects by a probabilistic observation of attribute values. Finally, in Section 6, the approach is extended to the case of a fuzzy attribute mapping, the attribute values of objects being described by means of possibility distributions of the attribute range, thus mixing rough sets and Dubois and Prade's twofold (ill-known) fuzzy sets [8].

2. Preliminaries and notation

Rough sets are induced by the impossibility to distinguish between objects in a data table by means of their precise description in terms of attribute values, because more attributes, or more refined attribute scales, would be needed to tell them apart. Then only upper and lower approximation of sets of objects defined in extension can be described by the available attributes. The opposite situation is when all objects have potentially distinct descriptions (attribute scales are sufficiently refined) but the attribute values are only partially known. Then, a set of objects defined by a certain property expressed in terms of attribute values can be ill-known in extension. This section recalls the two situations separately in more details, while the next section considers them jointly.

2.1. An attribute function generates a rough set

Let us establish some notation. Let $U = \{u_1, \ldots, u_n\}$ be a set of objects. Let $f : U \to V$ be an attribute function (it represents the observation of an attribute[1] for objects in U). $V = \{v_1, \ldots, v_m\}$ is the class of possible values of the attribute. Let $C_i = f^{-1}(\{v_i\})$ be the collection of objects described by value v_i, for each $i = 1, \ldots, m$. Clearly $\Pi = \{C_1, \ldots, C_m\}$ is a partition of U. For an arbitrary set of objects S we can define its upper and lower approximations by means of attribute f as:

$$\overline{\mathrm{appr}}_\Pi(S) = \cup\{C \in \Pi : C \cap S \neq \emptyset\}; \underline{\mathrm{appr}}_\Pi(S) = \cup\{C \in \Pi : C \subseteq S\}. \tag{1}$$

Thus S is an exact set when $\underline{\mathrm{appr}}_\Pi(S) = S = \overline{\mathrm{appr}}_\Pi(S)$. If not, it is called a *rough set*. In that case, we easily check that none of these equalities holds, i.e.,

$$\underline{\mathrm{appr}}_\Pi(S) \subsetneq S \subsetneq \overline{\mathrm{appr}}_\Pi(S).$$

It depicts the limited expressive power of attribute f for describing a set S defined in extension: assume that all we observe about each object $u \in U$ is the attribute value, $f(u)$. Then, all we know about the class S is that it contains $\underline{\mathrm{appr}}_\Pi(S)$ and it is contained in $\overline{\mathrm{appr}}_\Pi(S)$.

[1]Notice that V may be a Cartesian product of attribute domains, so f may represent a collection of attributes.

According to [18], rough sets and classification problems are related. For instance, let $d : U \to K$ be a decision function that classifies the objects of the universe U. For each $k \in K$, $S_k = d^{-1}(\{k\})$ represents a class of objects, and the family of classes $\{S_1, \ldots, S_k\}$ is another partition of U. The classification problem is one of approximating this partition by means of equivalence classes of the attribute mapping f. In other words, a class S_i is an exact set if and only if the observation of the attribute(s) suffices to classify all objects in or out it. Thus, all the classes are exacts sets when Π is a refinement of $\{S_1, \ldots, S_k\}$. Otherwise $\underline{\mathrm{appr}}_\Pi(S)$ and $\overline{\mathrm{appr}}_\Pi(S)$ represent all we know about the class S when we can only observe the attribute f.

2.2. An ill-known attribute function generates an ill-known set

Let a one-to-many mapping $F : U \to \wp(V)$ represent an imprecise observation of the attribute $f : U \to V$. Namely, for each object $u \in U$, all that is known about the attribute value $f(u)$ is that it belongs to the set $F(u) \subseteq V$. Suppose we want to describe the set $f^{-1}(A)$ of objects that satisfy a property $A \subset V$, namely $\{u \in U : f(u) \in A\}$. Because of the incompleteness of the information, the subset $f^{-1}(A) \subseteq U$ is an "ill-known set" [8]. Let us first recall the following definition:

Definition 2.1. ([7]) Let X and Y be two arbitrary sets and let $H : X \to \wp(Y)$ be a multi-valued mapping with non-empty images. Let $A \subseteq Y$ be an arbitrary subset of Y. The *upper inverse* of A is defined as

$$H^*(A) = \{x \in X \,:\, H(x) \cap A \neq \emptyset\}.$$

The *lower inverse* of A is defined as

$$H_*(A) = \{x \in X \,:\, H(x) \subseteq A\}.$$

According to this definition $f^{-1}(A)$ can be approximated from above and from below, respectively by upper and lower inverses of A via F:

- $F^*(A) = \{u \in U \,:\, F(u) \cap A \neq \emptyset\}$ is the set of objects that possibly belong to $f^{-1}(A)$.
- $F_*(A) = \{u \in U \,:\, F(u) \subseteq A\}$ is the set of objects that surely belong to $f^{-1}(A)$.

The Boolean algebra interval $[F_*(A), F^*(A)] = \{B, F_*(A) \subseteq B \subseteq F^*(A)\}$, called an interval set by Yao [28], contains the set $f^{-1}(A)$. The multi-valued mappings F^* and F_* are respectively upper and lower inverses of F.

3. Coverings induced by incomplete observations

There is a strong similarity between rough sets and ill-known sets. However the origin of uncertainty is radically different. In the case of rough sets, attribute values are known but there are not enough attributes (or attribute domains are too coarse) to single out objects by their descriptions using these attributes. In the case of ill-known sets there may be enough attribute values, but the lack of knowledge about objects forbids a precise enumeration of the contents of sets defined by means of properties. It is clear that these sources of uncertainty being unrelated, they can be simultaneously present. This section describes this hybrid situation. It is shown here that imprecise observations of the attribute can be described, in a natural way, by coverings. We are then led to study direct covering-based extensions of rough sets proposed by Pomykala and Yao already referred to in the introduction.

3.1. From ill-known attributes to coverings

Consider again the multimapping F between U and V. For each value $v \in V$, let us consider its upper inverse image, the subset of objects of U for which it is possible that $f(u) = v$:

$$F^*(\{v\}) = \{u \in U \,:\, v \in F(u)\} \subseteq U.$$

In other words, if $u \notin F^*(\{v_i\})$, we are sure that $f(u) \neq v_i$. Let $V = \{v_1, \ldots, v_m\}$ and $\mathcal{C} \subseteq \wp(U)$ be the following family of subsets of U:

$$\mathcal{C} = \{F^*(\{v\}) \,:\, v \in V\} = \{F^*(\{v_1\}), \ldots, F^*(\{v_m\})\}$$

Two important remarks must be made:

- The lower inverse image $F_*(\{v\})$ may easily be empty as it is the set of objects known to have attribute value v exactly. However, $F^*(\{v\}) = \emptyset$ only if no object may have value v. It may be also that $F^*(\{v\}) = U$, if all objects may have value v.
- Just like in the precisely-informed case of rough sets, the sets $F^*(\{v_i\})$ may not be distinct. It holds that $F^*(\{v_i\}) = F^*(\{v_j\}) = S$ if all objects in S may have value v_i or v_j (among other possible alternatives) and all objects outside S are known not to have these values. More generally, it is possible to define an equivalence relation $\sim_F$ on V defined as follows:

  $$v_i \sim_F v_j \iff F^*(\{v_i\}) = F^*(\{v_j\}).$$

  It defines a partition of V. Attribute values in an equivalence class V_k of $\sim_F$ play the same role in distinguishing objects. It is then useless to distinguish these values and it is convenient to replace V by the quotient set $V/\sim_F$, moreover deleting the class V_k such that $F^*(\{v\}) = \emptyset, \forall v \in V_k$ (i.e. attribute values taken by no object.)

These considerations allow us to assume that the attribute range V is such that $v_i \neq v_j$ implies $F^*(\{v_i\}) \neq F^*(\{v_j\})$, and $F^*(\{v\}) \neq \emptyset, \forall v \in V$. In other words the class $\mathcal{C} = \{F^*(\{v\}) \,:\, v \in V\} = \{F^*(\{v_1\}), \ldots, F^*(\{v_m\})\}$ contains distinct non-empty sets of objects. This assumption will be taken for granted in the remainder of the paper.

We will next show that $\mathcal{C}$ is a covering of U and that, under the above proviso, it basically determines the same information as the imprecise observation of the attribute, F, up to a renaming of attribute values $v \in V$.

Let us check it in the following theorem:

Theorem 3.1. Let $F : U \to \wp(V)$ be a multi-valued mapping defined on U.
Let $\mathcal{C} = \{F^*(\{v_1\}), \ldots, F^*(\{v_m\})\} = \{C_1, \ldots, C_m\}$. Then:

1. If $F(u) \neq \emptyset,\ \forall\, u \in U$ then $\mathcal{C}$ is a covering of U, i.e. $\cup_{i=1}^m C_i = U$.
2. The family of pairs $\{(C_1, v_1), \ldots, (C_m, v_m)\}$ univocally determines F.
3. The covering $\mathcal{C} = \{C_1, \ldots, C_m\}$ determines F only up to a possible permutation of elements of V.

Proof:

1. Let $u \in U$ be an arbitrary element in the universe. Since $F(u) \neq \emptyset$, there exists $i \in \{1, \ldots, m\}$ such that $v_i \in F(u)$, i.e. $u \in F^*(\{v_i\}) = C_i$. Hence, $u \in \cup_{i=1}^m C_i$.

2. By construction, $u \in F^*(\{v\}) \iff v \in F(u)$. If each element C_i of the covering is attached to an attribute value v_i, we can define $F(u) = \{v_i \,:\, u \in C_i\},\ \forall\, u \in U$. Hence, F is univocally determined by the family of pairs (C_i, v_i).

3. It is clear that the number of elements in the covering is $m \leq |V|$. By assumption, the attribute range V is such that $v_i \neq v_j$ implies $F^*(\{v_i\}) \neq F^*(\{v_j\})$, so, there is the same number of elements in V as in the covering.

So, $\mathcal{C}$ and F represent the same knowledge about the equivalence classes associated to the ill-known attribute function f modeled by F, the key-equivalence being $v \in F(u) \iff u \in F^*(\{v\})$. Note that, if only the covering $\mathcal{C}$, generated by F, is available, we can no longer retrieve the values v_i such that $F^*(\{v_i\}) = C_i$, hence the multi-mapping F is, strictly speaking, out of reach. However, we can recompute a set V' and an equivalent multi-mapping $F' : U \to \wp(V')$ (i.e. yielding the same covering), letting $V' = \{1, \ldots m\}$ and $F'(u) = \{i \in V' \,:\, u \in C_i\}$. Clearly V' has the same number of elements as $\mathcal{C}$ and V. In fact, the multi-mapping F can be recovered as soon as the bijection between $\mathcal{C}$ and V is known. The results in the last theorem are clarified in the following example.

Example 3.1. Let $U = \{u_1, u_2, u_3, u_4\}$. Let $V = \{v_1, v_2, v_3\}$ and let $F : U \to \wp(V)$ be defined as follows:

$$F(u_1) = \{v_1, v_2\}, \quad F(u_2) = \{v_1, v_3\}, \quad F(u_3) = \{v_2, v_3\}, \quad F(u_4) = \{v_3\}.$$

Let us suppose that F represents an imprecise information of an attribute $f : U \to V$. The covering associated to F, $\mathcal{C} = \{C_1, C_2, C_3\}$, is given by:

$$C_1 = F^*(\{v_1\}) = \{u_1, u_2\}, C_2 = F^*(\{v_2\}) = \{u_1, u_3\},$$
$$C_3 = F^*(\{v_3\}) = \{u_2, u_3, u_4\}.$$

Note that the multivalued mapping $F'(u_1) = \{1, 2\}, F'(u_2) = \{1, 3\}, F'(u_3) = \{2, 3\}, F'(u_4) = \{3\}$ also induces the same covering as F.

Let us suppose that we only know that $\mathcal{C} = \{C_1, C_2, C_3\}$ is the covering associated to a unknown multi-valued mapping F. In that case, we can assume that there are three different possible (useful) values $\{1, 2, 3\}$ for the attribute. F can then be retrieved as follows (up to a renaming of such values):

$$\begin{aligned} F(u_1) &= \{k \,:\, u_1 \in C_k\} = \{1, 2\} \\ F(u_2) &= \{k \,:\, u_2 \in C_k\} = \{1, 3\} \\ F(u_3) &= \{k \,:\, u_3 \in C_k\} = \{2, 3\} \\ F(u_4) &= \{k \,:\, u_4 \in C_k\} = \{3\} \end{aligned}$$

Remark 3.1. We easily check that $\mathcal{C}$ is a partition if and only if the images of F are disjoint. In that case, as noticed earlier at the beginning of this section, these images form a partition of V, and V can be replaced by its equivalence classes. Namely, we are back to the case when F represents the precise observation of an attribute f, (i.e., $F(u) = \{f(u)\},\ \forall\, u \in U$), and then $\mathcal{C}$ coincides with the partition $\Pi = \{f^{-1}(\{v_1\}), \ldots, f^{-1}(\{v_m\})\}$.

Remark 3.2. The covering can be a class of nested sets.

Let $V = \{v_1, \ldots, v_m\}$ be a finite set and let $C_i = F^*(\{v_i\})$. Then the covering $\mathcal{C} = \{C_1, \ldots, C_m\}$ is a family of nested sets if and only if the images of F are nested. In fact,

$$C_i \subseteq C_j \Leftrightarrow [v_i \in F(u) \Rightarrow v_j \in F(u)]$$

So, $C_1 \subseteq C_2 \subseteq \ldots \subseteq C_m$ if and only if the images of F are of the form $F(u) = \{v_j \,:\, j \geq i\} = \{v_i, \ldots, v_m\}$, for some $i \in \{1, \ldots, m\}$. So, the elements of the covering are nested if and only if the images of F are also nested. It means that the attribute may have the same value on all objects, but there are different degrees of precision in the perception of the attribute for each object of the population.

Remark 3.3. What happens when we observe l different attributes on the same universe of objects?

- Particular case: The images of F are Cartesian products. Let us suppose that $F(u) = F_1(u) \times \ldots \times F_l(u),\ \forall\, u \in U$. Then

$$\begin{aligned} F^*(\{(v_1, \ldots, v_l)\}) &= \{u \in U \,:\, (v_1, \ldots, v_l) \in F(u)\} \\ &= \{u \in U \,:\, v_1 \in F_1(u), \ldots, v_l \in F_l(u)\} = \cap_{i=1}^{l} F_i^*(\{v_i\}). \end{aligned}$$

  So, in this case, the elements of the covering $\mathcal{C}$ associated to F are the intersections of elements of the partial coverings $\mathcal{C}_i$. Furthermore, the binary relation R can be written as a function of the partial binary relations R_i, $i = 1, \ldots, l$. In fact, $uRu' \Leftrightarrow uR_iu', \forall\, i = 1, \ldots, l$. This case generalizes the case of precise observations of the l attributes.

- General case: The images of F are not necessarily Cartesian products. In this general case, we cannot ensure the equalities $F^*(\{(v_1, \ldots, v_l)\}) = \cap_{i=1}^{l} F_i^*(\{v_i\})$, because the inclusion

$$F^*(\{(v_1, \ldots, v_l)\}) \supseteq \cap_{i=1}^{l} F_i^*(\{v_i\})$$

  is not fulfilled in general. With respect to the binary relation, the implication $uRu' \Rightarrow uR_iu', \forall\, i = 1, \ldots, l$, holds, but the converse is not true in general. Let us give an example: let $U = \{u_1, u_2\}$ and let $l = 2$. let $V = V_1 \times V_2 = \{v_{11}, v_{12}\} \times \{v_{21}, v_{22}\}$. Let F be the multi-valued mapping defined as $F(u_1) = \{(v_{11}, v_{21}), (v_{12}, v_{22})\}$ and $F(u_2) = \{(v_{11}, v_{22}), (v_{12}, v_{21})\}$. Its respective projections on V_1 and V_2, F_1 and F_2, are vacuous, so they do not give any relevant information. In fact, $F_1(u_1) = F_1(u_2) = V_1 = \{v_{11}, v_{12}\}$ and $F_2(u_1) = F_2(u_2) = V_2 = \{v_{21}, v_{22}\}$. Hence, the corresponding binary relations R_1 and R_2 are also uninformative (u_1 and u_2 are R_1- and R_2-indistinguishable). Nevertheless, they are $R-$ distinguishable, since $F(u_1) \cap F(u_2) = \emptyset$. In other words, $u_1 \not R u_2$, but $u_1 R_1 u_2$ and $u_1 R_2 u_2$.

3.2. Coverings viewed as ill-known partitions

The multimapping F represents incomplete knowledge about the real attribute mapping $f : U \to V$. Each attribute mapping consistent with F generates a partition of U. As a consequence, the covering induced by F must be viewed as an ill-known partition. This will allow us to choose appropriate extensions of upper and lower approximations in the sense of a covering.

Definition 3.1. An attribute mapping f satisfying $f(u) \in F(u), \forall u \in U$, written $f \in F$ for short, is called a selection function of F.

Each such mapping f induces an equivalence relation R_f on U, namely $uR_f v \iff f(u) = f(v)$. The sets $\{f^{-1}(\{v\}) \neq \emptyset, v \in V\}$ form a partition Π_f of U. This partition enables f to be reconstructed (up to a renaming of elements in $f(U)$). The equivalence class of an element $u \in U$ with respect to R_f is denoted $[u]_f = f^{-1}(\{f(u)\})$. The upper (resp. lower) approximation of a subset S of objects according to attribute function f is $\overline{\text{appr}}_f(S) = \cup_{u \in S} f^{-1}(\{f(u)\})$ (resp. $\underline{\text{appr}}_f(S) = (\overline{\text{appr}}_f(S^c))^c$).

When the attribute mapping is ill-known and described by F, it generates a set of possible partitions Π_f of U, each associated with some attribute mapping $f \in F$. It is then natural to interpret the covering $\mathcal{C}$ induced by F as a set of possible partitions of U, one of which is the right one.

Proposition 3.1. The covering $\mathcal{C} = \{F^*(\{v\}) : v \in V\}$ induced by F is such that $\forall f \in F, \forall v \in V$, $\exists\, u \in U, [u]_f \subseteq F^*(\{v\})$. Moreover $F^*(\{v\}) = \cup_{f \in F} f^{-1}(\{v\})$.[2]

Proof:
Fix $f \in F$, and let u such that $f(u) = v$. Then $[u]_f = f^{-1}(\{v\})$. So $\forall u' \in [u]_f$, as $u' \in f^{-1}(\{v\})$, $f(u') = v \in F(u')$. Hence $u' \in F^*(\{v\})$. More generally, $F^*(\{v\}) = \{u \in U : v = f(u), f \in F\} = \cup_{f \in F}\{u \in U : v = f(u)\} = \cup_{f \in F} f^{-1}(\{v\})$.

So each element $F^*(\{v\})$ of the covering induced by F is the union of possible equivalence classes of objects whose attribute value is v. Or equivalently, if what is known is the covering $\mathcal{C} = \{C_1, \dots C_k\}$ interpreted as the set of partitions $\{\Pi = (E_1, \dots, E_k), E_j \subseteq C_j, j = 1, \dots k\}$ (where for convenience we admit that some E_j may be empty, and implicitly deleted). It consists in allocating each object $u \in U$ to a single set C_j among those to which it belongs in $\mathcal{C}$.

Example 3.2. Consider again Example 3.1, according to the information provided by F, such that $F(u_1) = \{v_1, v_2\}$, $F(u_2) = \{v_1, v_3\}$, $F(u_3) = \{v_2, v_3\}, F(u_4) = \{v_3\}$. It can be checked that F induces the following possible partitions of U, each obtained by choosing one element in $F(u_i), i = 1, \dots, m$:

[2] Remember that, by convention, $\forall v \in V, F^*(\{v\}) \neq \emptyset$.

$$\Pi_1 = \{\{u_1, u_2\}, \{u_3\}, \{u_4\}\}$$
$$\Pi_2 = \{\{u_1, u_2\}, \{u_3, u_4\}\}$$
$$\Pi_3 = \{\{u_1\}, \{u_3\}, \{u_2, u_4\}\}$$
$$\Pi_4 = \{\{u_1\}, \{u_2, u_3, u_4\}\}$$
$$\Pi_5 = \{\{u_1, u_3\}, \{u_2\}, \{u_4\}\}$$
$$\Pi_6 = \{\{u_1\}, \{u_2\}, \{u_3, u_4\}\}$$
$$\Pi_7 = \{\{u_1, u_3\}, \{u_2, u_4\}\}$$

We easily check that we can recover the class of possible partitions, $\{\Pi_1, \ldots, \Pi_7\}$, from $\mathcal{C}$ or equivalently, from F'. Namely, as $\mathcal{C} = \{\{u_1, u_2\}, \{u_1, u_3\}, \{u_2, u_3, u_4\}\}$, a partition is recovered by assigning to a single set each object u_i belonging to more than one element of the covering. So, u_1 can be re-assigned either to C_1 or C_2, u_2 either to C_1 or C_3, u_3 either to C_2 or C_3, u_4 can only belong to C_3. For instance, deciding to assign u_1 and u_2 to C_1, u_3 to C_3 yields Π_2, etc. Note that in this case, C_2 gets empty so that the obtained partition only has two elements.

More generally, the multimapping F corresponds to at most $\prod_{u \in U} |F(u)|$ partitions. This product is the number of selection functions of the multimapping F. In the example, F has 8 selections, but there are only 7 possible partitions, because Π_4 is obtained by two selection functions of F.

In the situation of incomplete information, a covering is interpreted differently from the case of, say, formal concept analysis [12], where elements of the covering are known concepts. Here, a covering should be interpreted as a set of possible partitions, one of which is the correct one. As seen above, these partitions are generated by all the selections of the multiple-valued attribute mapping representing imprecise attribute information. Each such selection is a precise attribute mapping generating a standard rough set. So, incomplete information mappings lead to ill-known rough sets.

4. From coverings to ill-known rough sets

In this section, we choose among the covering-based extensions of rough sets appearing in the literature, the ones that fit the setting of imprecise attribute mappings. We show that the relation between objects induced by an ill-known attribute function and its associated covering leads to the loose upper and lower approximations of a set proposed by Pomykala [24] and Yao[31], the loose upper approximation being formed by the union of covering elements that intersect the set. The mapping that assigns to each object its indistiguishability neighborhood is called the top-class mapping. The upper approximation of a set of objects is then the upper inverse image of this set via the top-class mapping. The lower approximation is defined by duality as the complement of the upper approximation of the complement of the set. We next lay bare the definition of a tight pair of approximations in that setting. Interestingly, the upper and lower approximations correspond respectively to all objects that surely belong to the rough upper approximation (a lower approximation of the rough upper approximation), and all objects that possibly belong to the rough lower approximation (an upper approximation of the rough lower approximation). But the tight pair of approximations does not seem to correspond to a proposal in the literature, although it can be approximated by another extension of rough sets proposed by Pomykala too.

4.1. Generalized rough sets based on coverings

Many generalizations of the concept of "rough set" can be found in the literature. One of them precisely consists in replacing the partition (induced by the equivalence relation) by a more general covering of the universe (see [24, 2, 29, 31, 32], for instance). See Samantha and Chakraborty [26] for an extensive classification of covering-based rough sets and their properties. The concepts of lower and upper approximations can be extended to this more general context in different ways. In fact there are several equivalent ways of defining the pairs of upper and lower approximations in the standard rough set setting, and these lead to different definitions in the case of coverings.

The most straightforward method is as follows [24]. Let $\mathcal{C}$ be a covering of the universe, U, i.e., a family of subsets of U such that $\cup\{C \in \mathcal{C}\} = U$. Let us replace the partition Π by $\mathcal{C}$ in Equation 1. In that case, the upper and lower approximations considered in Equation 1 are no longer dual. In fact, the class

$$\cup\{C \in \mathcal{C} \,:\, C \cap S \neq \emptyset\}$$

overlaps the set

$$\cup\{C \in \mathcal{C} \,:\, C \subseteq S^c\}.$$

hence does not coincide with its complement. This approach is followed by Yao ([31]) who considers the following two pairs of (dual) upper and lower approximations:

- What can be called the *loose pair*:

$$\begin{aligned} \overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) &= \cup\{C \in \mathcal{C} \,:\, C \cap S \neq \emptyset\} \\ \underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) &= [\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S^c)]^c = \{u \in U \,:\, \forall C \in \mathcal{C}\,[u \in C \Rightarrow C \subseteq S]\} \\ &= \cup\{C \in \mathcal{C} \,:\, C \subseteq S \wedge [\nexists C' \in \mathcal{C}, C' \cap S^c \neq \emptyset \wedge C \cap C' \neq \emptyset]\}. \end{aligned} \tag{2}$$

- What can be called the *tight pair* :

$$\begin{aligned} \underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) &= \cup\{C \in \mathcal{C} \,:\, C \subseteq S\} \\ \overline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) &= [\underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S^c)]^c = \{u \in U \,:\, \forall C \in \mathcal{C}\,, [u \in C \Rightarrow C \cap S \neq \emptyset]\}. \\ &= \cup\{C \in \mathcal{C} \,:\, C \cap S \neq \emptyset \wedge [\nexists C' \in \mathcal{C}, C' \subseteq S^c \wedge C \cap C' \neq \emptyset]\} \end{aligned} \tag{3}$$

It is clear that $\underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) \subset \underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S)$, generally, and that $\overline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) \subset \overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)$. So, by construction

$$\underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) \subset \underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) \subset S \subset \overline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) \subset \overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S),$$

which explains the names "loose and tight" approximation pairs.

There are other possible ways of extending rough sets with coverings. For instance, Bonikowski *et al.* [2] rely on the duality between intensions (properties) and extensions (sets of objects) along the line of formal concept analysis[12]. Then, a covering is a set of known concepts or properties [3]. So, an object is described by the elements of the coverings it belongs to. The minimal description $M(u)$ of object u is the set of minimal elements C of the coverings that contain u. Then two objects u and u' are

[3]Not exactly what we call attributes, even if you can define a property as a Boolean attribute

indiscernible if they have the same minimal description. One is led to considering the Boolean algebra induced by $\mathcal{C}$ on U, whose atoms are made of indiscernible objects.

Bonikowski *et al.* then define the lower approximation of a subset S of objects as $\underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S)$, but he defines the upper approximation as $\underline{\mathrm{appr}}^{B}_{\mathcal{C}}(S) = \underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) \cup Bn(S)$, where $Bn(S)$ is the boundary of S obtained as

$$Bn(S) = \cup_{u \in S \setminus \underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S)} \cup_{C_i \in M(u)} C_i$$

That is, considering the boundary as unions of neighborhoods (obtained from $M(u)$) of elements in S that are not in its lower approximation. This boundary is very thick and overlaps the lower approximation.

In the next section, we show that the approximations of sets of objects that arise in a natural way from imprecise observations of an attribute or a set thereof form the two pairs in Equation (2) proposed by Pomykala and studied by Yao.

4.2. The top-class function and the associated tolerance relation

It is possible to define a kind of neighborhood around each object $u \in U$ by considering the upper inverse image of its ill-known attribute value:

Definition 4.1. Let us consider the multi-valued mapping $F : U \to \wp(V)$. Let us define another multi-valued mapping $I\!F : U \to \wp(U)$ as follows:

$$I\!F(u) := F^*(F(u)) = \{u' \in U \,:\, F(u') \cap F(u) \neq \emptyset\},\ \forall\, u \in U.$$

$I\!F$ is called the *top-class function* associated to F.

We may interpret the top-class $I\!F(u)$ as the neighborhood of u. It is the set of elements that could belong to the equivalence class $f^{-1}(\{f(u)\})$ of objects not distinguishable from u, induced by the attribute f whose imprecise description is F, namely:

1. If the observation of the attribute is precise (i.e., $F(u) = \{f(u)\},\ \forall\, u \in U$), the top-class function is of the form $I\!F(u) = f^{-1}(\{f(u)\}) = \{u' \in U \,:\, f(u') = f(u)\}$. In words, $I\!F$ associates, to each object $u \in U$, its equivalence class (the class of objects that are indistinguishable from u.) (If R is the equivalence relation associated to F, then $I\!F(u) = [u]_R = \{u' \in U \,:\, uRu'\},\ \forall\, u \in U$.)

2. Let us now suppose that F represents an imprecise observation of the attribute $f : U \to V$. The above multi-valued function, $I\!F : U \to \wp(U)$, associates, to each $u \in U$, the most precise set that contains its equivalence class. In words, all that is known about the objects in U that are possibly indistinguishable from u is that they belong to $I\!F(u)$. Another way to put it is that if $u' \notin I\!F(u)$, then u' and u can not belong to the same equivalence class.

The following lemma is instrumental in the proof of Theorem 4.1.

Lemma 4.1. Let $F : U \to \wp(V)$ be the multi-valued function that represents the imprecise observation of the attribute and let $\mathcal{C}$ be the covering associated to F. Let $I\!F : U \to \wp(U)$ be the top-class function associated to F. Then

$$\forall\, u \in U, I\!F(u) = \cup\{C \in \mathcal{C} \,:\, u \in C\} = \bigcup_{v \in F(u)} F^*(\{v\}).$$

Proof:
We easily check that $u' \in I\!F(u)$ if and only if there exists $C \in \mathcal{C}$ such that $u, u' \in C$. Indeed consider C where $u \in C$. Then $\exists\, v_i \in V, u \in F^*(v_i) = C$ which means $v_i \in F(u)$ exactly. This attribute value v_i does not depend on the choice of $u \in C$. If moreover $u' \in C$ too, $v_i \in F(u')$ as well. Hence $F(u) \cap F(u') \neq \emptyset$. Hence $u' \in I\!F(u)$, so that $C \subseteq I\!F(u)$. Conversely, if $u' \in I\!F(u)$ then, by construction, $\exists\, v_i \in V, u, u' \in F^*(v_i) = C \in \mathcal{C}$. In other words, $u' \in I\!F(u)$ if and only if $u' \in \cup\{C \in \mathcal{C} \,:\, u \in C\}$. Besides, it is clear that $\{C \in \mathcal{C} \,:\, u \in C\} = \{F^*(\{v\}) \,:\, v \in F(u)\}$.

Remark 4.1. The top-class $I\!F(u)$ clearly differs from the minimal description $M(u)$ of Bonikowski induced by the covering, as the former involves no minimality requirement. In fact $M(u) \subset I\!F(u)$. The neighborhood of u defined by $I\!F(u)$ was also proposed by Pomykala [24] and is often denoted $Friends(u)$. Kryszkiewicz [15] also lays bare this notion, for a special case of ill-known attribute function whereby attribute values are either known or totally unknown.

The top-class function, $I\!F$, does not univocally determine the covering $\mathcal{C}$ as seen in the following example.

Example 4.1. Let $U = \{u_1, u_2, u_3\}$, $V = \{1, 2, 3\}$ and let $F : U \to \wp(V)$ be defined as follows: $F(u_1) = \{1, 2\},\ F(u_2) = \{2, 3\},\ F(u_3) = \{1, 3\}$. The top-class function, $I\!F : U \to \wp(U)$, associated to F is given by:

$$I\!F(u_1) = I\!F(u_2) = I\!F(u_3) = \{u_1, u_2, u_3\}.$$

Let us now consider another multi-valued mapping, $F' : U \to \wp(V)$, defined as $F'(u_i) = \{1, 2, 3\},\ \forall\, i = 1, 2, 3$. We easily check that the top-class function associated to F' coincides with $I\!F$. The information provided by F is more precise than the information provided by F', since $F(u_i) \subsetneq F'(u_i),\ i = 1, 2, 3$. Nevertheless, the top-class functions associated to F and F' do coincide; in particular, $I\!F$ is trivial and has clearly lost all the information contained in F.

Interestingly, we have built a multiple-valued mapping $I\!F$ within U, and we can naturally consider upper and lower inverses of subsets of U in the sense of Dempster. The following theorem shows that they coincide with the upper approximations of sets of objects induced by the ill-known attribute mapping by collecting elements of the covering that intersect the given set, i.e., Pomykala-Yao's loose pair of upper and lower approximations (Equation 2) coincide with the upper and lower inverse images of $I\!F$.

Theorem 4.1. Let us consider a multi-valued mapping $F : U \to \wp(V)$. Let $\mathcal{C} = \{F^*(\{v\}) \,:\, v \in V\} = \{C_1, \ldots, C_m\}$ the covering of U associated to F. Let $S \subseteq U$ be an arbitrary subset of U. Let $I\!F : U \to \wp(U)$ be the top-class function associated to F. Let $\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)$ and $\underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)$ be the upper and lower approximations of S of Equation 2. Then:

$$\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) = I\!F^*(S) \text{ and } \underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) = I\!F_*(S).$$

Proof:
According to Lemma 4.1, $I\!F(u) = \cup\{C \in \mathcal{C} \,:\, u \in C\}$. Hence:

$$I\!F^*(S) = \{u \in U \,:\, I\!F(u) \cap S \neq \emptyset\} =$$

$$\{u \in U \,:\, \cup\{C \in \mathcal{C} \,:\, u \in C\} \cap S \neq \emptyset\} = \cup\{C \in \mathcal{C} \,:\, C \cap S \neq \emptyset\}.$$

This class coincides (see Equation 2) with Pomykala-Yao's first upper approximation $\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)$. On the other hand, Dempster upper and lower images are dual, and thus:

$$I\!F_*(S) = [I\!F^*(S^c)]^c = [\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S^c)]^c = \underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S).$$

This pair of approximations is also used by Kryszkiewicz[15] in his special framework for incomplete knowledge.

Instead of the top-class function, we can lay bare a relation among objects induced by the covering, and retrieve these upper and lower approximations by applying the rough set definitions to this relation.

Definition 4.2. F induces the following binary relation $R_F \subseteq U \times U$: $u\, R_F u' \Leftrightarrow F(u) \cap F(u') \neq \emptyset$.

In other words: $u\, R_F u' \Leftrightarrow \exists C \in \mathcal{C}$ such that $u, u' \in C$. We will call R_F the *binary relation induced by the multi-valued mapping F* (or *the binary relation induced by the covering $\mathcal{C}$*). We easily check the following:

Proposition 4.1. Relation R_F is reflexive and symmetric, but it is not transitive, in general. Furthermore, the following assertions are equivalent:

- R_F is transitive
- $C_i, C_j \in \mathcal{C} \Rightarrow C_i = C_j$ or $C_i \cap C_j = \emptyset$.
- The different images of F are disjoint, i.e, $[F(u) \cap F(u') \neq \emptyset \Rightarrow F(u) = F(u')]$.

Orlowska & Pawlak [20] interpret uR_Fu' as a similarity between u and u', and indeed these two properties are characteristic of the mathematical notion of similarity. But this terminology may be misleading in our context since it suggests a measure of proximity between elements while uR_Fu' only represents *potential* similarity between u and u'. For instance, if $F(u) = F(u') = V$, then the attribute values of u and u' are totally unknown, and so is their actual similarity, while uR_Fu' holds.

According to definition 4.2, $u\, R_F\, u'$ if and only if these elements *may* be in the same equivalence class.[4] In other words, they are not R_F-related when we are sure that they do not belong to the same equivalence class, that is, we are sure that $f(u) \neq f(u'), \forall f \in F$, i.e., they are not associated to the same attribute value. [5]

The above binary relation R_F is univocally determined by the top-class function $I\!F$. In fact, R_F can be obtained from $I\!F$ as follows:

$$u' R_F u \Leftrightarrow u' \in I\!F(u) (\Leftrightarrow u \in I\!F(u')).$$

Conversely, $I\!F$ is determined by R_F as follows:

$$I\!F(u) = \{u' \in U \,:\, u' R_F u\}, \forall\, u \in U.$$

[4]Here we refer to the equivalence classes of equivalence relation associated to the true values of the attribute. This equivalence relation is not known. We only know that it is some subset of R_F.

[5]Kryszkiewicz [15] also uses this relation in his restricted framework for incompleteness, with a similar comment.

Let us emphasize again that the covering $\mathcal{C}$ provides more information than any of R_F or $I\!F$. If uR_Fu', you only know that their images by F do overlap. So, you know that there is at least some C in the covering such that $u, u' \in C$. But you do not know how many nor which $C's$ satisfy this condition. In fact, two different coverings may induce the same binary relation, since they can be associated to same two-class function, as example 4.1 shows.

From now on, let us use the following notation:

$$[u]_{R_F} = I\!F(u) = \{u' \in U \,:\, uR_Fu'\} = \{u' \in U \,:\, F(u) \cap F(u') \neq \emptyset\}$$

According to this notation, the upper and lower approximations of S can be alternatively written just like for standard rough sets:

$$\begin{array}{rcl} \overline{\mathrm{appr^L}}_{\mathcal{C}}(S) & = & \{u \in U \,:\, [u]_{R_F} \cap S \neq \emptyset\} \\ \underline{\mathrm{appr^L}}_{\mathcal{C}}(S) & = & \{u \in U \,:\, [u]_{R_F} \subseteq S\} \end{array} \tag{4}$$

Remark 4.2. Yao and Lingras say in [30] that the concepts of upper and lower approximations can be extended as in Equation 4, when R is a general binary relation. But they do not try to relate this type of formulation (Equation 4) with the upper and lower approximations for coverings given in Equations 2 and 3.

Contrary to the precise standard rough set case, the set $F^*(A) = \{u \in U \,:\, F(u) \cap A \neq \emptyset\}$ of objects whose attribute value possibly lies in a subset A of V is not representable in terms of the associated covering. Namely, $\overline{\mathrm{appr^L}}_{\mathcal{C}}(F^*(A))$ does not necessarily coincide with $F^*(A)$. Indeed, for all $A \in 2^V$,

$$\begin{array}{rcl} \overline{\mathrm{appr^L}}_{\mathcal{C}}(F^*(A)) & = & \{u \in U \,:\, [u]_{R_F} \cap F^*(A) \neq \emptyset\} \\ & = & \{u \in U \,:\, \{u' : F(u) \cap F(u') \neq \emptyset\} \cap \{u'', F(u'') \cap A \neq \emptyset\} \neq \emptyset\} \\ & = & \{u \in U \,:\, \exists\, u' \text{ with } F(u) \cap F(u') \neq \emptyset \text{ and } F(u') \cap A \neq \emptyset\} \end{array}$$

However the above set only contains $\{u \in U \,:\, F(u) \cap A \neq \emptyset\} = F^*(A)$ (this inclusion is easy to check just enforcing $u = u'$). Let, for instance, F be defined on $U = \{u_1, u_2\}$ as follows: $F(u_1) = \{1, 2\}$ and $F(u_2) = \{2, 3\}$ and let $A = \{3\}$. Then $F^*(A) = \{u_2\}$ but u_1 belongs to $\overline{\mathrm{appr^L}}_{\mathcal{C}}(F^*(A))$, because $F(u_1) \cap F(u_2) \neq \emptyset$ and $F(u_2) \cap A \neq \emptyset$ (although $F(u_1) \cap A = \emptyset$.)

This fact contrasts with the deterministic case where a deterministic mapping f is used, and where $F^*(A) = f^{-1}(A)$ is an exact set, with respect to the equivalence relation on U induced by f.

Remark 4.3. Contrary to the standard rough set case, the family of loose upper approximations $\{I\!F^*(A) \,:\, A \subseteq U\} = \{\overline{\mathrm{appr^L}}_{\mathcal{C}}(A) \,:\, A \subseteq U\}$ does not determine the multimapping F, nor the covering $\mathcal{C}$ (we checked it in Example 4.1). It could be of interest to look for the "reduct" of the cover $\mathcal{C}$, i.e., the least precise covering associated to the same $I\!F$.

4.3. The possible world approach to upper and lower approximations of rough sets

The peculiar behavior of upper inverse images $F^*(A)$ as not closed under representation by a covering leads us to claim that a covering induced by a multimapping associated to an ill-known attribute does not play the same role as a partition in the theory of rough sets. Namely, we have to really consider the covering as a family of possible partitions induced by the selection functions $f \in F$. Each partition represents a possible situation characterized by a precise description of attribute values, hence a possible actual database. This is the possible world approach [17].

Under this view, the upper approximation $\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)$ in the sense of the covering as induced by the top-class function turns out to be the union of possible standard rough set upper approximations of S induced by selections of F:

Proposition 4.2. $\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) = \cup_{f\in F}\overline{\mathrm{appr}}_f(S)$; $\underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) = \cap_{f\in F}\underline{\mathrm{appr}}_f(S)$.

Proof:
$\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) = \cup_{v\in V}\{\cup_{f\in F} f^{-1}(\{v\}) : \cup_{f\in F} f^{-1}(\{v\}) \cap S \neq \emptyset\}$ using Proposition 3.1. Factorizing $\cup_{f\in F}$ out yields:
$\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) = \cup_{f\in F}\cup_{v\in V}\{f^{-1}(\{v\}) : f^{-1}(\{v\}) \cap S \neq \emptyset\} = \cup_{f\in F}\overline{\mathrm{appr}}_f(S)$.

The other result obtains by duality: $\underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) = (\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S^c))^c$
$= [\cup_{f\in F}\overline{\mathrm{appr}}_f(S^c)]^c = \cap_{f\in F}(\overline{\mathrm{appr}}_f(S^c))^c = \cap_{f\in F}\underline{\mathrm{appr}}_f(S)$.

So a given V-related coverage $\mathcal{C}$ such that $C_i = F^*(\{v_i\})$, we can always select in each C_i a (possibly empty) single subset [6] D_i so that the non-empty $D_i's$ forms a partition Π. Each such partition defines an attribute mapping f such that $\forall i, \forall u \in D_i, f(u) = v_i$. So a covering can be viewed as a Boolean possibility distribution over partitions. Note that from the above results, if the only information about each element $u \in U$ is the imprecise measurement $F(u)$ and S is a particular class of objects described by some property, then, it is known that S includes $\underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)$ and is included in $\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)$.

However more information is available, namely, all objects $u \in \cap_{f\in F}\overline{\mathrm{appr}}_f(S)$ do belong to the upper approximation of S induced by the ill-known attribute function, and $\cup_{f\in F}\underline{\mathrm{appr}}_f(S)$ contains the ill-known lower approximation of S. The following proposition studies the relationship between these sets and the tight pairs of covering-based rough sets.

Proposition 4.3. $\underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) \subseteq \cup_{f\in F}\underline{\mathrm{appr}}_f(S)$; $\overline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) \supseteq \cap_{f\in F}\overline{\mathrm{appr}}_f(S)$.

Proof:
Note that the alternative lower approximation due to Pomykala-Yao: $\underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) = \cup\{C \in \mathcal{C} : C \subseteq S\}$ can be written $\cup_{v\in V}\{\cup_{f\in F} f^{-1}(\{v\}) : \cup_{f\in F} f^{-1}(\{v\}) \subseteq S\} \subseteq \cup_{f\in F}\cup_{v\in V}\{f^{-1}(\{v\}) : f^{-1}(\{v\}) \subseteq S\} = \cup_{f\in F}\underline{\mathrm{appr}}_f(S)$. The other expression follows by duality.

The converse inclusions in the above proposition may fail to hold.

Example 4.2. Suppose U and V only have two elements. Let $F(u_1) = V$ and $F(u_2) = \{v_2\}$. There are only two possible attribute mappings: $f_1(u_1) = v_2 = f_1(u_2)$; $f_2(u_1) = v_1$ and $f_2(u_2) = v_2$. Consider approximating $S = \{u_2\}$. Then $f_1^{-1}(v_2) = U \not\subset S$, while $f_1^{-1}(v_1) = \emptyset$, and $f_2^{-1}v_1) = \{u_1\}, f_2^{-1}(v_2) = \{u_2\}$. So $\cup_{f\in F}\underline{\mathrm{appr}}_f(S) = S$, but clearly $f_1^{-1}(\{v_1\}) \cup f_1^{-1}(\{v_2\}) = U \not\subset S$ and $f_2^{-1}(\{v_1\}) \cup f_2^{-1}(\{v_2\}) = U \not\subset S$. Hence $\underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) = \emptyset$. So the tight pair is not useful in approximating the upper and lower rough approximation of S.

[6] We assume the finite subsets $F^*(\{v_i\})$ are distinct if not empty.

So the above construct gives a natural interpretation to definitions of loose upper and lower approximations with respect to a covering following Pomykala and Yao, not of the tight approximations. In summary:

$$\begin{aligned}\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) &= \cup\{C \in \mathcal{C} : C \cap S \neq \emptyset\} = \cup_{f \in F}\overline{\mathrm{appr}}_f(S) \\ \underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S) &= (\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S^c))^c = \cap_{f \in F}\underline{\mathrm{appr}}_f(S).\end{aligned} \tag{5}$$

and

$$\begin{aligned}\underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) &= \cup\{C \in \mathcal{C} : C \subseteq S\} \subseteq \cup_{f \in F}\underline{\mathrm{appr}}_f(S) \\ \overline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S) &= [\underline{\mathrm{appr}^{\mathrm{T}}}_{\mathcal{C}}(S^c)]^c \supseteq \cap_{f \in F}\overline{\mathrm{appr}}_f(S).\end{aligned} \tag{6}$$

The ill-known upper approximation of S is bracketed by the pair $(\cup_{f \in F}\overline{\mathrm{appr}}_f(S), \underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S))$ and the ill-known lower approximation of S is bracketed by the pair $(\underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S), \cap_{f \in F}\overline{\mathrm{appr}}_f(S)$, laying bare the fact that there is a double source of uncertainty about S: the one due to indiscernibility related to the attribute mapping, and the one due to the imprecision about this attribute mapping. One open question is now to find a simple lower approximation of $\cap_{f \in F}\overline{\mathrm{appr}}_f(S)$ and upper approximation of $\cup_{f \in F}\underline{\mathrm{appr}}_f(S)$, preferably among the many notions of covering-based rough sets. It is not clear that the tight upper and lower approximations defined via the selection functions can be expressed in terms of the covering induced by the incomplete attribute mapping.

5. Generalized accuracy measures

The concept of “rough probability” proposed by Pawlak can be extended to this more general context of reasoning with imprecisely known objects. This notion tries to measure the accuracy of rough sets.

5.1. Rough probability

Following Pawlak [21] the upper and lower *qualities* of S are defined as

$$\overline{q}(S) = \frac{|\overline{\mathrm{appr}}_\Pi(S)|}{|U|} \qquad \text{and} \qquad \underline{q}(S) = \frac{|\underline{\mathrm{appr}}_\Pi(S)|}{|U|} \tag{7}$$

The difference $\overline{q}(S) - \underline{q}(S)$ reflects how well S can be described by the partition Π. Pawlak also considers the accuracy of approximation of S by the partition as the quantity $\alpha_R(S) = \frac{\underline{q}(S)}{\overline{q}(S)}$. Note that from a mathematical point of view, it is well known that $\underline{q}$ and $\overline{q}$ are respectively belief and plausibility measures *with respect to mass function defined on the partition by* $m(C_i) = \frac{|C_i|}{|U|}$. However, this terminology is not correct in the sense that the rough set setting does not consider there exists an ill-known object to be discovered. What is ill-known here is the set S of objects. So, it sounds better to consider these evaluations as inner and outer measures of S, due to the fact that S is not measurable in the language of the partition (as done by Fagin and Halpern[11]).

Pawlak[22] also suggested generalizing these concepts by starting from a probability function, P, on the $\sigma-$ algebra $\sigma(\Pi)$ (instead of using the cardinality of sets), and defining the lower and upper probabilities of S as:

$$\overline{P}_\Pi(S) = \inf P(\{C : C \in \sigma(\Pi),\ C \supseteq S\}) = P(\overline{\mathrm{appr}}_\Pi(S)),$$

$$\underline{P}_\Pi(S) = \sup P(\{C \,:\, C \in \sigma(\Pi),\ C \subseteq S\}) = P(\underline{\mathrm{appr}}_\Pi(S)).$$

The probability measure on U corresponds to a way of assigning importance weights to elements (e.g. they can be duplicated), so that importance weights affect quality functions and accuracy evaluations. When, in particular, P is the Laplace probability, i.e., when $P(C) = \frac{|C|}{|U|},\ \forall C \in \sigma(\Pi)$, then $\underline{P}_\Pi$ and $\overline{P}_\Pi$ respectively coincide with the lower and upper qualities of S. Pawlak [22] referred to the interval $[\underline{P}_\Pi(S), \overline{P}_\Pi(S)]$ as the *rough probability* of S. The same comment as above applies here, if we consider (U, Π, P) as a probability space: we can only compute upper and lower probabilities of subsets not because there is some ill-known element in U described by an imprecise probability function, but because we cannot describe all subsets properly with the partition.

However we may suggest the following interpretation of rough probabilities: Assume that all we know about each object $u \in U$ is the value of the attribute on it $f(u)$. Suppose that, for each $i \in \{1, \ldots, m\}$, the probability of selecting an object from $C_i = f^{-1}(\{v_i\})$ is known. Then, all we know about the probability of selecting an object from the class S is that it lies between $\underline{P}_\Pi(S)$ and $\overline{P}_\Pi(S)$.

5.2. Accuracy of an ill-known rough set

Since the conjoint occurrence of indiscernibility and measurement imprecision of the attribute leads to introduce imprecision in the knowledge of the upper and lower approximations of a subset of objects, we get imprecise evaluations $\overline{q}_{\mathcal{C}}(S), \underline{q}_{\mathcal{C}}(S)$ of the quality functions (7) due to the lack of knowledge of the attribute mapping $f \in F$ inducing a covering $\mathcal{C}$. Namely

$$\overline{q}_{\mathcal{C}}(S) = [\frac{|\cap_{f\in F} \overline{\mathrm{appr}}_f(S)|}{|U|}, \frac{|\cup_{f\in F} \overline{\mathrm{appr}}_f(S)|}{|U|}]$$

and

$$\underline{q}_{\mathcal{C}}(S) = [\frac{|\cap_{f\in F} \underline{\mathrm{appr}}_f(S)|}{|U|}, \frac{|\cup_{f\in F} \underline{\mathrm{appr}}_f(S)|}{|U|}]$$

Our results enable an efficient computation of the upper bound of $\overline{q}_{\mathcal{C}}(S)$ as $\frac{|\overline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)|}{|U|}$ and of the lower bound of $\underline{q}_{\mathcal{C}}(S)$ as $\frac{|\underline{\mathrm{appr}^{\mathrm{L}}}_{\mathcal{C}}(S)|}{|U|}$.

The accuracy of approximation of S by the ill-known $f \in F$ is the interval

$$\tilde{\alpha}_R(S) = [\inf_{f\in F} \frac{\underline{q}_f(S)}{\overline{q}_f(S)}, \sup_{f\in F} \frac{\underline{q}_f(S)}{\overline{q}_f(S)}]$$

This quantity cannot be easily computed by separate calculations on the numerator and the denominator because they may not use to the same selection function f in numerator and denominator, as requested by the definition.

Imprecise rough membership function : The Laplacean probability $P(S|u)$ that an object u belongs to S is only known to lie in interval

$$[\inf_{f\in F} \frac{|[u]_f \cap S|}{|[u]_f|}, \sup_{f\in F} \frac{|[u]_f \cap S|}{|[u]_f|}]$$

We could generalize these evaluations to when a probability distribution exists on U, like in the case of rough probabilities. In this case we get bracketings of the inner and outer probabilities $P(\overline{\mathrm{appr}}_f(S))$ and $P(\underline{\mathrm{appr}}_f(S))$.

5.3. Imprecise“rough” probability

Let us first recall the notions of Dempster upper and lower probabilities.

Definition 5.1. ([7]) Let $H : X \to \wp(Y)$ be a multi-valued mapping with non-empty images. Let σ_X be a $\sigma-$algebra on X. Let $P : \sigma_X \to [0,1]$ be a probability measure. Let $A \subseteq Y$ be an arbitrary subset of Y such that $H^*(A) \in \sigma_X$. The *upper probability* of A is defined as

$$P^*(A) = P(H^*(A)) = P(\{x \in X \,:\, H(x) \cap A \neq \emptyset\}).$$

The *lower probability* of A is defined as

$$P_*(A) = P(H_*(A)) = P(\{x \in X \,:\, H(x) \subseteq A\}).$$

Dempster lower probabilities are $\infty-$order monotone capacities. Thus, when Y is a finite universe, they are a pair of belief and plausibility functions in the sense of Shafer[27]. We can extend the concept of “rough probability” to the case of coverings as follows.

Definition 5.2. Let $F : U \to \wp(V)$ an arbitrary multi-valued mapping defined on the finite universe U and let $\mathcal{C}$ be the associated covering. Let P be a probability measure induced by a probability distribution on U. Let us define $\overline{P}_\mathcal{C} : \wp(U) \to [0,1]$ and $\underline{P}_\mathcal{C} : \wp(U) \to [0,1]$ as follows:

$$\overline{P}_\mathcal{C}(S) := P(\overline{\mathrm{appr}^{\mathrm{L}}}_\mathcal{C}(S)), \quad \underline{P}_\mathcal{C}(S) := P(\underline{\mathrm{appr}^{\mathrm{L}}}_\mathcal{C}(S)),\ \forall\, S \subseteq U.$$

Call $[\underline{P}_\mathcal{C}(S), \overline{P}_\mathcal{C}(S)]$ the *covering-rough probability* of S.

Note that one may be tempted to define P on the σ-algebra generated by $\mathcal{C}$. But this algebra makes little sense in our setting as the intersection of elements of a covering do not generally contain any equivalence class induced by the ill-known attribute mapping f.

Theorem 5.1. Let $\overline{P}_\mathcal{C} : \wp(U) \to [0,1]$ and $\underline{P}_\mathcal{C} : \wp(U) \to [0,1]$ denote the covering-upper and lower probabilities in the previous definition. Then, $\overline{P}_\mathcal{C}$ and $\underline{P}_\mathcal{C}$ are respectively a plausibility and a belief function.

Proof:
Notice that they coincide with the Dempster upper and lower probabilities induced by the multi-valued mapping $I\!F$, i.e.,

$$\overline{P}_\mathcal{C}(S) = P(\overline{\mathrm{appr}^{\mathrm{L}}}_\mathcal{C}(S)) = P(I\!F^*(S)) = P^*(S) \text{ and}$$

$$\underline{P}_\mathcal{C}(S) = P(\underline{\mathrm{appr}^{\mathrm{L}}}_\mathcal{C}(S)) = P(I\!F_*(S)) = P_*(S),\ \forall\, S \subseteq U.$$

When $\mathcal{C}$ is a partition, $[\underline{P}_\mathcal{C}(S), \overline{P}_\mathcal{C}(S)]$ clearly coincides Pawlak’s rough probability, $\forall\, S \subseteq U$. The same reservations apply to the interpretation obtained set-functions as belief and plausibility functions:

they are actually upper and lower bounds of weighted quality functions $P(\overline{\text{appr}}_f(S))$ and $P(\underline{\text{appr}}_f(S))$, respectively. Moreover, the lower bound of $P(\overline{\text{appr}}_f(S))$, and the upper bound of $P(\underline{\text{appr}}_f(S))$ can be obtained likewise, but it is not clear that they are infinite monotone or alternating functions.

Like in the complete information case the covering-rough probability could nevertheless be interpreted as follows: Assume that, for each object $u \in U$, we can only observe an imprecise measurement, $F(u)$, of a certain characteristic, f. On the other hand, let S be a class of objects (in a certain classification process $d : U \to K$.) Under this information, all I know about the probability of selecting an element from the class S is that it is between $\underline{P}_{\mathcal{C}}(S)$ and $\overline{P}_{\mathcal{C}}(S)$. One may assume that the probability values on the sets $C \in \mathcal{C}$ are known (or they can be estimated from a sample of the population), so $\underline{P}_{\mathcal{C}}(S)$ and $\overline{P}_{\mathcal{C}}(S)$ can be derived (or estimated from a sample).

Remark 5.1. Another view is that V is equipped with a probability distribution p_V (values $f(u)$ were observed for not directly accessible objects u). In the latter case, we consider indiscernible the outcomes u that yield the same $f(u)$, and measurable the subsets of U exactly described via f. Then, we can define a mass function m_f on U, letting $m_f(f^{-1}(\{v\})) = p_V(v), \forall v \in V$. However, the fact that we only know that $f \in F$ prevents us from assigning the mass $p_V(v)$ to a precise subset of U. Then we may assign it to $F^*(v)$ as being the largest subset containing the real equivalence class. It defines a random set on U with mass function $\overline{m}$ such that each random set $m_f, f \in F$, is a specialization of $\overline{m}$. Conversely we could assign the masses $p_V(v)$ to $\cap_{f\in F} f^{-1}(\{v\})$. We would then get a random set on U with mass function $\underline{m}$ that is a specialization of each random set $m_f, f \in F$.

6. Fuzzy rough sets induced by fuzzy attribute mappings : preliminary steps

In this section we show how to extend the previous construction when the ill-known attribute mapping is represented by means of a fuzzy mapping. Ill-known sets are then represented by two nested fuzzy sets called twofold fuzzy sets[8]. On the other hand, fuzzy rough sets [9] have been constructed by changing the equivalence relation into a similarity relation. This section bridges the gap between the two notions, by first constructing the fuzzy relation induced by the fuzzy attribute mapping and then building the fuzzy loose approximations as a variant of fuzzy rough set.

6.1. Twofold fuzzy sets

From now on, let $\mathcal{F}(X)$ represent the class of fuzzy subsets of an arbitrary universe X. Let $\tilde{F} : U \to \mathcal{F}(V)$ represent an imprecise observation of the attribute $f : U \to V$, i.e., for each $u \in U$, $\tilde{F}(u)$ is a possibility distribution over V that represents our imprecise knowledge about $f(u)$:

> For an arbitrary pair $u \in U$, $v \in V$, the possibility degree that v coincides with the true attribute value $f(u)$ is $\tilde{F}(u)(v) \in [0, 1]$.

In particular we assume that $\exists v \in V, \tilde{F}(u)(v) = 1$, which extends to the fuzzy case the requirement that $F(u) \neq \emptyset$. For each $\alpha \in (0, 1]$, let us now denote by $\tilde{F}_\alpha : U \to \wp(V)$ the $\alpha-$ cut of $\tilde{F}$, i.e., the multimapping:

$$\tilde{F}_\alpha(u) = \{ v \in V : \tilde{F}(u)(v) \geq \alpha\}. \tag{8}$$

According to [10, 3], this multimapping induces the following interpretation of $\tilde{F}$ is equivalent to the following one:

> For each $\alpha \in (0,1]$, the probability that $f(u)$ belongs to $\tilde{F}_\alpha(u),\ \forall\, u \in U$, is greater or equal to $1-\alpha$.

Dubois and Prade [8] considered the problem of describing the set $f^{-1}(A) \subseteq U$ of objects that satisfy a crisp property A, when f is only known via $\tilde{F}$. Because of incomplete information, the subset $f^{-1}(A)$ is an *ill-known set* bracketed by a pair of fuzzy sets $\tilde{F}^*(A)$ and $\tilde{F}_*(A)$ defined by:

- $\mu_{\tilde{F}^*(A)} = \sup_{v \in A} \mu_{\tilde{F}(u)}(v)$: all objects that are more or less possibly in $f^{-1}(A)$.
- $\mu_{\tilde{F}_*(A)} = \inf_{v \notin A} 1 - \mu_{\tilde{F}(u)}(v)$: all objects that more or less surely belong to $f^{-1}(A)$.

The pair $(\tilde{F}_*(A), \tilde{F}^*(A))$ is such that Support($\tilde{F}_*(A)$) $\subseteq f^{-1}(A) \subseteq$ core($\tilde{F}^*(A)$) and was called a twofold fuzzy set.

6.2. Fuzzy rough sets

A fuzzy relation $\tilde{R}$ is a fuzzy set of $U \times U$. When symmetric and reflexive and min-transitive ($\tilde{R}(u_1,u_3) \geq \min(\tilde{R}(x_1,u_2)\tilde{R}(u_2,u_3))$, it is called a similarity relation by Zadeh [33]. It is a straightforward generalization of an equivalence relation, as the α-cuts of $\tilde{R}$ are nested equivalence relations. The value $R(u_1,u_2)$ evaluates the *actual* similarity between precise objects, contrary to Orlowska-Pawlak definition of similarity [20]. Indeed, Zadeh's similarity measures are closely related to metrics and ultrametrics ($1 - R(u_1,u_2)$ is a generalized form of metric). Definitions of rough sets can be extended to similarity relations:

Definition 6.1. [9] Let $\tilde{R} : U \times U \to [0,1]$ be a similarity relation on U. Let F be a fuzzy subset of U. The lower approximation of F is the fuzzy set:

$$\underline{\mathrm{appr}}_{\tilde{R}}(F)(u) = \inf_{u' \in U} \max\{1 - \tilde{R}(u,u'), F(u')\}$$

The upper approximation of F is the fuzzy set:

$$\overline{\mathrm{appr}}_{\tilde{R}}(F)(u) = \sup_{u' \in U} \min\{\tilde{R}(u,u'), F(u')\}$$

When $\tilde{R}$ is reflexive and symmetric, it is clear that the support of $\tilde{R}$ yields a covering $\mathcal{C}_{\tilde{R}} = \{\{u', \tilde{R}(u,u') > 0\}, u \in U\}$. Denote $C^i_{\tilde{R}} = \{u', \tilde{R}(u_i,u') > 0\}, 1, \ldots, |U|$. The supports of $\underline{\mathrm{appr}}_{\tilde{R}}(F)$ and $\overline{\mathrm{appr}}_{\tilde{R}}(F)$ are respectively equal to $\underline{\mathrm{appr}}^T_{\mathcal{C}_{\tilde{\mathcal{R}}}}(S)$ and $\overline{\mathrm{appr}}^L_{\mathcal{C}_{\tilde{\mathcal{R}}}}(S)$ if S is the support of F.

Note that other definitions have been proposed later on, where the min and max operators are replaced by general $T-$ norms and conorms, and the implication $\max(1-a,b)$ at work in the definition of upper approximation is replaced by a residuated implication [25]. The above definition preserves the duality between upper and lower approximations while in general this is not true. Moreover, $\underline{\mathrm{appr}}_{\tilde{R}}(F)(u) = 1$ means that the fuzzy neighborhood $\tilde{R}(\cdot,u)$ of u entirely lies in the core of F, which enforces a very strong view of fuzzy set-inclusion.

Noticing that a fuzzy similarity relation $\tilde{R}$ (reflexive, symmetric and transitive in the sense of a t-norm [25]) is such that its similarity classes$\{\tilde{R}(u,\cdot), u \in U\}$ overlap, they form a fuzzy covering of U. For this reason, De Cock *et al.* [5] have proposed several other definitions of upper and lower approximations of fuzzy sets especially the following ones (their notation) :

$$\tilde{R} \downarrow\downarrow F(u) = \inf_{u' \in U} \mathcal{I}(\tilde{R}(u,u'), \inf_{u'' \in U} \mathcal{I}(\tilde{R}(u',u''), F(u''))),$$

$$\tilde{R} \downarrow\uparrow F(u) = \inf_{u' \in U} \mathcal{I}(\tilde{R}(u,u'), \sup_{u'' \in U} \mathcal{T}(\tilde{R}(u',u''), F(u''))).$$

where $\mathcal{I}$ is an implication operation, and $\mathcal{T}$ a t-norm. Using the covering induced by the support of the fuzzy relation $\tilde{R}$, $\tilde{R} \downarrow\downarrow S$ and $\tilde{R} \downarrow\uparrow S$ have supports respectively defined by

$$(\tilde{R} \downarrow\downarrow S)_0 = \{u, \forall i = 1, \dots, |U|, u \in C^i_{\tilde{R}} \implies C^i_{\tilde{R}} \subseteq S\} = \underline{\mathrm{appr}}^L_{\mathcal{C}_{\tilde{\mathcal{R}}}}(S),$$

$$(\tilde{R} \downarrow\uparrow S)_0 = \{u, \forall i = 1, \dots, |U|, u \in C^i_{\tilde{R}} \implies C^i_{\tilde{R}} \cap S \neq \emptyset\} = \overline{\mathrm{appr}}^T_{\mathcal{C}_{\tilde{\mathcal{R}}}}(S).$$

The first is a loose lower approximation in the sense of Eq. (2), and the second is a tight upper approximation in the sense of Eq. (3). In this paper, we will only deal with the fuzzy-upper and lower approximations of crisp sets $S \subseteq U$. According to Definition 6.1, when, in particular, we approximate a crisp set, S, we have:

$$\begin{array}{rcl} \overline{\mathrm{appr}}_{\tilde{R}}(S)(u) & = & \sup_{u' \in S} \tilde{R}(u,u'). \\ \underline{\mathrm{appr}}_{\tilde{R}}(S)(u) & = & \inf_{u' \notin S}[1 - \tilde{R}(u,u')]. \end{array} \tag{9}$$

For each α, the α-cut of the upper and lower approximations of S coincides with the upper and lower approximations of S for the α-cut of $\tilde{R}$, the (classical) equivalence relation $R_\alpha \subseteq U \times U$ defined by $u \, R_\alpha \, u' \Leftrightarrow \tilde{R}(u,u') \geq \alpha$. Again, Support($\underline{\mathrm{appr}}_{\tilde{R}}(S)) \subseteq S \subseteq$ core($\overline{\mathrm{appr}}_{\tilde{R}}(S)$). The above fuzzy rough set is in the spirit of the loose approximation pairs of Pomykala and Yao's definition of covering-based approximations, changing the tolerance relation induced by the covering into the fuzzy relation $\tilde{R}$.

6.3. Fuzzy relation associated to a fuzzy–valued imprecise observation of the attribute

For each $\alpha \in (0,1]$, the multi-valued mapping $\tilde{F}_\alpha : U \to \wp(V)$ from Equation 8 plays the same role as F in Section 3. Thus, for each $\alpha \in (0,1]$, we can derive a binary reflexive and symmetric relation R_α as follows:

Definition 6.2. Let $\tilde{F} : U \to \mathcal{F}(V)$ be a fuzzy-valued mapping from U to V and let $\tilde{F}_\alpha : U \to \wp(V)$ its $\alpha-$cut, $\forall \, \alpha \in (0,1]$. Let $R_\alpha \subseteq U \times U$ be defined as follows:

$$u \; R_\alpha u' \Leftrightarrow \tilde{F}_\alpha(u) \cap \tilde{F}_\alpha(u') \neq \emptyset.$$

This is the $\alpha-$ *level relation associated to* $\tilde{F}$. We can then derive a binary fuzzy relation from $\tilde{F}$ in a natural way as follows:

Definition 6.3. Let $\tilde{F} : U \to \mathcal{F}(V)$ be a fuzzy-valued mapping from U to V. For each $\alpha \in (0,1]$, let R_α be the $\alpha-$level binary relation associated to $\tilde{F}$. Let $R_{\tilde{F}} : U \times U \to [0,1]$ be defined as follows:

$$R_{\tilde{F}}(u,u') = \sup\{\alpha \in (0,1] \; : \; uR_\alpha u'\} = \sup\{\alpha \in (0,1] \; : \; \tilde{F}_\alpha(u) \cap \tilde{F}_\alpha(u') \neq \emptyset\}.$$

This is the fuzzy relation associated to $\tilde{F}$. Equivalently it can be directly defined by

$$R_{\tilde{F}}(u,u') = \sup_{v\in V} \min(\mu_{\tilde{F}(u)}(v), \mu_{\tilde{F}(u')}(v))$$

Proposition 6.1. R is reflexive and symmetric, but it is not min–transitive in general. The following statements are equivalent:

1. R is min-transitive
2. R_α is an equivalence relation for each $\alpha \in (0,1]$
3. The images of $\tilde{F}_\alpha$ are disjoint for each $\alpha \in (0,1]$, i.e., $\tilde{F}_\alpha(u) \cap \tilde{F}_\alpha(u') \neq \emptyset \Rightarrow \tilde{F}_\alpha(u) = \tilde{F}_\alpha(u')$.

$R_{\tilde{F}}(x,y)$ can be interpreted as the degree of possibility of the following assertion: "x and y belong to the same equivalence class".

6.4. Fuzzy rough set derived from a fuzzy attribute mapping

Thanks to the fuzzy relation $R_{\tilde{F}}$, a fuzzy–valued imprecise attribute function $\tilde{F}$ induces fuzzy upper and lower approximations in a natural way, just applying the definition of fuzzy rough sets.Namely, the formulae given in Equation 9 express loose fuzzy upper and lower approximations of an ill-known crisp rough set bracketing a crisp set S, under imprecise observation of the attribute represented by $\tilde{F} : U \to \mathcal{F}(V)$.

Let S be an arbitrary crisp subset of U and let us define the *upper and lower fuzzy–approximations of S*, denoted by $\overline{\text{appr}}_{R_{\tilde{F}}}(S)$ and $\underline{\text{appr}}_{R_{\tilde{F}}}(S)$, as follows:

$$\overline{\text{appr}}_{R_{\tilde{F}}}(S)(u) = \sup_{u'\in S} R_{\tilde{F}}(u,u'),\ \forall\, u \in U, \tag{10}$$
$$\underline{\text{appr}}_{R_{\tilde{F}}}(S)(u) = \inf_{u'\notin S} [1 - R_{\tilde{F}}(u,u')],\ \forall\, u \in U. \tag{11}$$

Under our present interpretation of fuzzy sets, $\overline{\text{appr}}_{R_{\tilde{F}}}(S)(u)$ still represents the degree of possibility that u belongs to S and $\underline{\text{appr}}_{R_{\tilde{F}}}(S)(u)$ represents the degree of necessity (security) that u belongs to S.

There is a connection between the above fuzzy-upper and lower approximations and the loose upper and lower approximations of the cuts of the fuzzy attribute multimapping. Consider $\mathcal{C}_\alpha = \{\tilde{F}^*_\alpha(v), :\ v \in V\}$ the covering induced by $\tilde{F}_\alpha$. In fact, we easily check that the fuzzy upper and lower approximations (10) can be calculated as follows from $\overline{\text{appr}}_{\mathcal{C}_\alpha}(S)$ and $\underline{\text{appr}}_{\mathcal{C}_\alpha}(S)$:

$$\overline{\text{appr}}_{R_{\tilde{F}}}(S)(u) = \sup\{\alpha \in (0,1]\ :\ u \in \overline{\text{appr}}^L_{\mathcal{C}_\alpha}(S)\}$$

$$\underline{\text{appr}}_{R_{\tilde{F}}}(S)(u) = \sup\{\alpha \in (0,1]\ :\ u \in \underline{\text{appr}}^L_{\mathcal{C}_\alpha}(S)\},$$

using the loose upper and lower approximations of Pomykala and Yao wrt covering $\mathcal{C}_\alpha$. Using the interpretation of possibility degree as an upper probability bound, each α-cut of the fuzzy rough set $(\underline{\text{appr}}_{R_{\tilde{F}}}(S), \overline{\text{appr}}_{R_{\tilde{F}}}(S))$ is a pair of sets bracketing S with probability at least $1-\alpha$.

Remark 6.1. Note that the fuzzy mapping $\tilde{F}$ induces the fuzzy covering $\tilde{\mathcal{C}} = \{\tilde{C}_v, v \in V\}$, with membership grades such that $\tilde{C}_v(u) = \mu_{\tilde{F}(u)}(v)$. Then it can be checked that, for the fuzzy loose lower approximation:

$$\underline{\mathrm{appr}}_{R_{\tilde{F}}}(S)(u) = \inf_{v \in V} \mathcal{I}(\tilde{C}_v(u), \inf_{u' \in U} \mathcal{I}(\tilde{C}_v(u'), S(u'))) = \tilde{R}_{\tilde{F}} \downarrow\downarrow S(u),$$

with $S(\cdot)$ the characteristic function of set S, and $\mathcal{I}(a, b) = \max(1 - a, b)$.

Proof:
$\underline{\mathrm{appr}}_{R_{\tilde{F}}}(S)(u) = \inf_{u' \notin S} 1 - R_{\tilde{F}}(u, u') = \inf_{u' \notin S} 1 - \sup_{v \in V} \min(\tilde{C}_v(u), \tilde{C}_v(u'))$
$= \inf_{u' \notin S} \inf_{v \in V} \max(1 - \tilde{C}_v(u), 1 - \tilde{C}_v(u')) = \inf_{v \in V} \max(1 - \tilde{C}_v(u), \inf_{u' \notin S} 1 - \tilde{C}_v(u'))$
$= \inf_{v \in V} \mathcal{I}(\tilde{C}_v(u), \inf_{u' \in U} \mathcal{I}(\tilde{C}_v(u'), S(u')))$.

So, the fuzzy loose lower approximation induced by an uncertain mapping provides a suitable motivation for one of the original fuzzy lower rough approximations proposed by De Cock *et al.* [5]. However, it justifies the use of a non-residuated implication, contrary to what they suggest.

7. Conclusion and perspectives

This paper proposes an interpretation of covering-based rough sets as modelling an ill-known rough set induced both by the ill-observation of attribute values and the lack of discrimination of the set of attributes. In this setting where two sources of uncertainty are present, it is shown that one upper and one lower approximation is not enough: in fact the rough approximations (in the rough set sense) are themselves bracketed from above and from below since ill-known. In this context, the covering is viewed as an ill-known partition, each partition generating a possible family of rough sets. The loose upper (resp. lower) approximation in the sense of Pomikala and Yao, obtained by applying the usual upper approximation definition to the covering instead of the partition elements (resp. its dual via complementation), is shown to make sense in this setting and contains (resp. is contained in) all possible upper (resp. lower) approximations of a set in the sense of all such partitions.

The rough set construction based on the tolerance relation induced by the covering (associated to the ill-known attribute mapping) only justifies the loose approximations. So, this choice of covering-based generalization of rough sets is justified by the incomplete information semantics. Tight approximations have been defined, which together with the loose ones bracket the ill-known upper and lower rough approximations; but they seem to be difficult to compute and they do not seem to belong to the catalogue of existing covering-based rough sets. Unfortunately, the companion pairs of covering-based definitions, made by replacing the partition by the covering in the usual definition of lower approximation cannot serve as a tight inner approximation to the ill-known rough set. There seems to be more work to be done to get a simple expression of lower (resp. upper) approximation of the upper (resp. lower) rough approximation, that we rigorously defined via a possible world approach.

This work is a first step to relate incomplete knowledge and covering-based rough sets. A further step maybe to interpret connectives like conjunction, disjunction and implication of covering-based rough sets within the incomplete information setting, using the implication lattices laid bare by Samanta and Chakraborty [26]. Our approach also opens several perspectives. There is a need to reinterpret the numerous works done about incomplete information databases in the light of our framework, especially works by Nakata and colleagues [19], but also special cases like [15]. The section on fuzzy mappings should also be completed in order to recover tight fuzzy approximations and a more detailed comparison with the framework of De Cock *et al.* [5] would be fruitful. This step requires considering a fuzzy mapping as a possibility distribution over selection functions as done in the setting of fuzzy random variables in the sense of Kwakernaak[16], Kruse and Meyer[14]. It does not look difficult to study fuzzy extensions of fuzzy quality indices and fuzzy rough probabilities. Nevertheless, they may be instrumental to bridge the gap between rough sets and imprecise probabilities. Finally, other kinds of interpretive frameworks for covering-based rough sets deserve consideration, such as formal concept analysis.

Acknowledgements

The first author acknowledges partial support by FEDER funds and Spanish Grants TIN2008-06681-C06-04 and TIN2010-20900-C04-01. The second author acknowledges support from Université Paul Sabatier in the form of an invited professorship, which made this joint work possible. Both authors thanks the referees for numerous comments, including pointing out a defect in a result, thus helping us to significantly improve the presentation and content.

References

[1] C. Baudrit, I. Couso, D. Dubois, Joint propagation of probability and possibility in risk analysis: towards a formal framework, International Journal of Approximate Reasoning, 45, 82-105 (2007).

[2] Z. Bonikowski, E. Bryniarski, U. Wybraniec-Skardowska, Extensions and intentions in the rough set theory, Information Sciences, 107, 149-167 (1998).

[3] I. Couso, P. Gil, S. Montes, The necessity of the strong α-cuts of a fuzzy set, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 9, 249-262 (2001).

[4] I. Couso, L. Sánchez, Higher order models for fuzzy random variables, Fuzzy Sets and Systems, 159, 237-258 (2008).

[5] M. De Cock, C. Cornelis, E.E. Kerre, Fuzzy rough sets, the forgotten connection, IEEE trans. Fuzzy Systems, 15: 121-129 (2007).

[6] T. Deng, Y. Chen, W. Xu and Q. Dai, A novel approach to fuzzy rough sets based on a fuzzy covering, Information Sciences, 177: 2308-2326 (2007).

[7] A.P. Dempster, Upper and Lower Probabilities Induced by a Multivalued Mapping, The Annals of Mathematical Statistics, 38 (1967), pp. 325-339.

[8] D. Dubois, H. Prade, Twofold fuzzy sets and rough sets - Some issues in knowledge representation, Fuzzy Sets and Systems, 23, 3-18 (1987).

[9] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems, 17, 191-209 (1990).

[10] D. Dubois, H. Prade, When upper probabilities are possibility measures, Fuzzy Sets and Systems, 49, 65-74 (1992).

[11] R. Fagin, J. Y. Halpern, Uncertainty, belief, and probability. Computational Intelligence 7: 160-173 (1991).

[12] B. Ganter, R. Wille, Formal Concept Analysis, Springer-Verlag, 1999

[13] J. W. Grzymala-Busse, On the Unknown Attribute Values in Learning from Examples. Methodologies for Intelligent Systems (ISMIS 1991), Lecture Notes in Computer Science, vol. 542, Springer, Berlin, 1991, pp 368-377.

[14] R. Kruse, K.D. Meyer, Statistics with Vague Data, Reidel, Dordrecht, 1987.

[15] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences, 112, 39-49 (1998).

[16] H. Kwakernaak, Fuzzy random variables I, II, Information Sciences, 15 (1978) 129 17 (1979) 253278.

[17] W. Lipski, On databases with incomplete information, Journal of the ACM 28 : 41 - 70 (1981)

[18] M. Magnani, Rough Set Theory for Knowlege Discovery in Data Bases, Technical Report, 2003.

[19] M. Nakata, H. Sakai, Lower and Upper Approximations in Data Tables Containing Possibilistic Information, J.F. Peters *et al.* (eds), Transactions on Rough Sets VII, LNCS 4400, 170-189 (2007).

[20] E. Orlowska, Z. Pawlak: Representation of Nondeterministic Information. Theor. Comput. Sci. 29: 27-39 (1984)

[21] Z. Pawlak, Rough classification, International Journal of Man-Machine Studies 20, 469–483 (1984).

[22] Z. Pawlak, Rough probability, Bulletin of Polish Academy of Sciences. Mathematics 32, 607–615 (1984).

[23] Z. Pawlak, Rough Sets, Kluwer Academic Publisher, Dordrecht, 1991.

[24] J.A. Pomykala, Approximation, Similarity and Rough Constructions, ILLC Pre-publication Series for Computation and Complexity Theory, CT-93-07,University of Amsterdam.

[25] A. M. Radzikowska, E. E. Kerre A comparative study of fuzzy rough sets Fuzzy Sets and Systems,126: 137-155 (2002).

[26] P. Samanta, M. K. Chakraborty: Covering Based Approaches to Rough Sets and Implication Lattices. In : Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC 2009): Lecture Notes in Computer Science, vol. 5908, Springer, 2009, pp. 127-134.

[27] G. Shafer (1976) A mathematical theory of evidence. Princeton university press Princeton, NJ

[28] Y. Y. Yao: Interval-Set Algebra for Qualitative Knowledge Representation. Int. Conf. Computing and Information (1993) 370-374

[29] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information Sciences, 111, 239-259, 1998.

[30] Y.Y. Yao. and P.J. Lingras, Interpretations of belief functions in the theory of rough sets, Information Sciences, 104, 81-106 (1998).

[31] Y.Y. Yao, On generalizing rough set theory, In : Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Wang, G. Liu, Q. Yao, Y. Skowron, A. (Eds.), Lecture Notes in Computer Science, Springer, Vol. 2639, 44-51 (2003).

[32] W.H. Xu and W.X. Zhang, Measuring roughness of generalized rough sets induced by a covering, Fuzzy Sets and Systems, 158, 2443 - 2455 (2007).

[33] L.A. Zadeh, Similarity relations and fuzzy orderings, Information Sciences, 177-200 (1971).