Representative Set of Objects in Rough Sets Based on Galois Connections

Nicolás Madrid[1] and Eloísa Ramírez-Poussa[2(✉)]

[1] Department of Applied Mathematics, Universidad de Málaga, Málaga, Spain
`nicolas.madrid@uma.es`
[2] Department of Mathematics, Universidad de Cádiz, Cádiz, Spain
`eloisa.ramirez@uca.es`

Abstract. This paper introduces a novel definition, called representative set of objects of a decision class, in the framework of decision systems based on rough sets. The idea behind such a notion is to consider subsets of objects that characterize the different classes given by a decision system. Besides the formal definition of representative set of objects of a decision class, we present different mathematical properties of such sets and a relationship with classification tasks based on rough sets.

Keywords: Rough Set Theory · Attribute reduction · Object reduction reduct · Decision systems

1 Introduction

Rough Set Theory (RST) is a mathematical theory that have shown its suitability for practical tasks [11,24]. The search to increase its range of application has given rise to different generalizations of this theory [6,7,10,28] as well as the relationships with other theories [4,18,25].

There exist different procedures to define the basic operators of rough sets, i.e., the lower and upper approximation, as those based on element operators, granular classes or subsystems [26]. In this work, we consider the approximation operators given by interior and closure operators obtained from the composition of operators in an isotone Galois connection [3,8], which has been built from a slight modification of the operators in [27]. This idea has been already considered in other works, as [14,17,20,21], and has two important advantages:

– the use of the operators introduced in [27] could lead us to the following situation: the lower approximation of a set may not be contained in the set and its upper approximation may not contain the set. The consideration of interior and closure operators avoids such a situation.
– the approximation operators obtained from the interior and closure operator are more accurate that those in [27] (see [17]).

In this paper we focus on the reduction of objects for a classification task. This kind of reduction has been seldom considered by the research community, which has mainly focused on attribute reduction [2,5,12,18]. Some examples of the study of object reductions are [15], which analyses the reduction of objects oriented to keep the original attribute reducts and [1,13,16,22,23] that reducts objects and attributes in parallel. The present paper is oriented in a different way than the existing approaches dealing with object reduction. We show that when the indiscernibility relation is not an equivalence relation, the objects in the different classes of a classification task can be characterized by only few objects in the class; we call that objects representative of the decision class. In such a way, the representative objects of the decision classes can be used as clusters in classification tasks. In this paper, we provide the formal definition of the set of representative objects of a decision class and analyze its mathematical properties.

The paper is organized as follows: Sect. 2 introduces the definitions of the approximation operators based on isotone Galois connections considered in this work, together with some results needed to understand this work. In Sect. 3, we present the formal definition of the set of representative objects of a decision class and analyze its mathematical properties. Section 4 provides the conclusions and presents some prospect for future work.

2   Preliminaries

In this section we recall some basic notions in order to make the contribution as self-contained as possible.

The first notion we have to recall is the notion of approximation space.

Definition 1. *An* approximation space *is a pair (U, R), where U is a set (called universe) and R is a binary relation over U.*

In this work, we consider approximation spaces whose relation R can be an arbitrary relation. This fact leads us to distinguish between left and right relationships, and to generalize the standard definition of R-foreset.

Definition 2. *Let (U, R) be an approximation space, the sets defined as:*

$$xR = \{y \in U \mid (x, y) \in R\} \ and \ Ry = \{x \in U \mid (x, y) \in R\}$$

are the R-right-foreset of $x \in U$ and the R-left-foreset of $y \in U$, respectively.

From the previous generalization, four different approximation operators arises.

Definition 3. *Let (U, R) be an approximation space and $A \subseteq U$. We define the following operators:*

- $R {\downarrow}^r A = \{x \in U \mid xR \subseteq A\}$
- $R {\uparrow}_r A = \{x \in U \mid xR \cap A \neq \varnothing\}$
- $R {\downarrow}^\ell A = \{y \in U \mid Ry \subseteq A\}$
- $R {\uparrow}_\ell A = \{y \in U \mid Ry \cap A \neq \varnothing\}.$

It is important to highlight that the approximation operators $R {\downarrow}^\ell$ and $R {\uparrow}_\ell$ coincide with those presented in [27]. Additionally, the equalities $R {\downarrow}^r = R {\downarrow}^\ell = R {\downarrow}$ and $R {\uparrow}_r = R {\uparrow}_\ell = R {\uparrow}$ are satisfied, when the relation is symmetric. For such a reason, hereafter, if R is symmetric, we will write $R {\downarrow}$ and $R {\uparrow}$ instead of $R {\downarrow}^r$, $R {\downarrow}^\ell$ and $R {\uparrow}_r$, $R {\uparrow}_\ell$, respectively.

On the other hand, the pairs $(R {\uparrow}_r, R {\downarrow}^\ell)$ and $(R {\uparrow}_\ell, R {\downarrow}^r)$ are isotone Galois connections [8,9], whose definition is recalled below.

Definition 4. *Let $(P, \leq_P)$ and $(Q, \leq_Q)$ be posets. A pair (φ, ψ) of mappings $\varphi \colon P \to Q$, $\psi \colon Q \to P$ is called* isotone Galois connection *between P and Q if the following equivalence is satisfied, for all $p \in P$ and $q \in Q$:*

$$\varphi(p) \leq_Q q \quad \textit{if and only if} \quad p \leq_P \psi(q).$$

This notion is also called adjunction*. The mapping φ is called* lower (or left) adjoint *of ψ and the mapping ψ* upper (or right) adjoint *of φ.*

At this point, it is important to point out that in the case of considering arbitrary relations, the operators $R {\downarrow}_r$ and $R {\downarrow}_\ell$ may be unsuitable to represent lower approximations, since the inequalities $R {\downarrow}_r (A) \subseteq A$ or $R {\downarrow}_\ell (A) \subseteq A$ may not hold for some set $A \subseteq U$. Similarly, $R {\uparrow}_r$ and $R {\uparrow}_\ell$ may be unsuitable to represent upper approximations, since $A \subseteq R {\uparrow}_r (A)$ or $A \subseteq R {\uparrow}_\ell (A)$ could not be satisfied for some set $A \subseteq U$. However, the compositions $R {\uparrow}_r (R {\downarrow}^\ell (A))$ and $R {\uparrow}_\ell (R {\downarrow}^r (A))$ are always contained in A, whereas A is always contained in the composition $R {\downarrow}^\ell (R {\uparrow}_r (A))$ and $R {\downarrow}^r (R {\uparrow}_\ell (A))$.

Certainly, the inequalities $R {\downarrow}_r (A) \subseteq A$, $R {\downarrow}_\ell (A) \subseteq A$, $A \subseteq R {\uparrow}_r (A)$ and $A \subseteq R {\uparrow}_\ell (A)$ are satisfied for reflexive relations. But even in that case, the composition of these operators provide better approximations than considering simply $R {\downarrow}_r$, $R {\downarrow}_\ell$, $R {\uparrow}_r$ and $R {\uparrow}_\ell$. That is, if R is reflexive, we have

$$R {\downarrow}^\ell (A) \subseteq R {\uparrow}_r (R {\downarrow}^\ell (A)) \subseteq A \subseteq R {\downarrow}^\ell (R {\uparrow}_r (A)) \subseteq R {\uparrow}_r (A).$$

and

$$R {\downarrow}^r (A) \subseteq R {\uparrow}_\ell (R {\downarrow}^r (A)) \subseteq A \subseteq R {\downarrow}^r (R {\uparrow}_\ell (A)) \subseteq R {\uparrow}_\ell (A).$$

for all $A \subseteq U$.

In such a way, the notion of rough set is defined for arbitrary relations by using the following definition.

Definition 5. *Let (U, R) be an approximation space and $A \subseteq U$. The* lower *approximations of A are defined as:*

$$R{\uparrow}_r \left(R{\downarrow}^{\ell} (A) \right) \quad and \quad R{\uparrow}_{\ell} \left(R{\downarrow}^r (A) \right)$$

and the upper *approximations of A are defined as:*

$$R{\downarrow}^{\ell} \left(R{\uparrow}_r (A) \right) \quad and \quad R{\downarrow}^r \left(R{\uparrow}_{\ell} (A) \right).$$

A set $A \subseteq U$ is called a generalized rough set *if it is different from the two lower approximations and from the two upper approximations.*

The following theorem summarizes some basic properties of such compositions.

Theorem 1. *Let (U, R) be an approximation space and $A, B \subseteq U$, then:*

- *If $A \subseteq B$ then $R{\uparrow}_r \left(R{\downarrow}^{\ell} (A) \right) \subseteq R{\uparrow}_r \left(R{\downarrow}^{\ell} (B) \right)$*
- *If $A \subseteq B$ then $R{\uparrow}_{\ell} \left(R{\downarrow}^r (A) \right) \subseteq R{\uparrow}_{\ell} \left(R{\downarrow}^r (B) \right)$*
- *If $A \subseteq B$ then $R{\downarrow}^{\ell} \left(R{\uparrow}_r (A) \right) \subseteq R{\downarrow}^{\ell} \left(R{\uparrow}_r (B) \right)$*
- *If $A \subseteq B$ then $R{\downarrow}^r \left(R{\uparrow}_{\ell} (A) \right) \subseteq R{\downarrow}^r \left(R{\uparrow}_{\ell} (B) \right)$*
- *$R{\uparrow}_r \left(R{\downarrow}^{\ell} (A) \right) \subseteq A \subseteq R{\downarrow}^r \left(R{\uparrow}_{\ell} (A) \right)$*
- *$R{\uparrow}_r \left(R{\downarrow}^{\ell} (A) \right) \subseteq A \subseteq R{\downarrow}^{\ell} \left(R{\uparrow}_r (A) \right)$*
- *$R{\uparrow}_{\ell} \left(R{\downarrow}^r (A) \right) \subseteq A \subseteq R{\downarrow}^{\ell} \left(R{\uparrow}_r (A) \right)$*
- *$R{\uparrow}_{\ell} \left(R{\downarrow}^r (A) \right) \subseteq A \subseteq R{\downarrow}^r \left(R{\uparrow}_{\ell} (A) \right)$*
- *$R{\uparrow}_r \left(R{\downarrow}^{\ell} \left(R{\uparrow}_r \left(R{\downarrow}^{\ell} (A) \right) \right) \right) = R{\uparrow}_r \left(R{\downarrow}^{\ell} (A) \right)$*
- *$R{\uparrow}_{\ell} \left(R{\downarrow}^r \left(R{\uparrow}_{\ell} \left(R{\downarrow}^r (A) \right) \right) \right) = R{\uparrow}_{\ell} \left(R{\downarrow}^r (A) \right)$*
- *$R{\downarrow}^{\ell} \left(R{\uparrow}_r \left(R{\downarrow}^{\ell} \left(R{\uparrow}_r (A) \right) \right) \right) = R{\downarrow}^{\ell} \left(R{\uparrow}_r (A) \right)$*
- *$R{\downarrow}^r \left(R{\uparrow}_{\ell} \left(R{\downarrow}^r \left(R{\uparrow}_{\ell} (A) \right) \right) \right) = R{\downarrow}^r \left(R{\uparrow}_{\ell} (A) \right)$.*

In [19] is stated that the approximation operators in Definition 5 coincide with those of [27] when the relation is a preorder.

Theorem 2 [19, Theorem 1]. *Let (U, R) be an approximation space. The following items are equivalent:*

- *$R{\uparrow}_{\ell} \left(R{\downarrow}^r (A) \right) = R{\downarrow}^r (A)$, for all $A \subseteq U$.*
- *$R{\uparrow}_r \left(R{\downarrow}^{\ell} (A) \right) = R{\downarrow}^{\ell} (A)$, for all $A \subseteq U$.*
- *R is a preorder (i.e. R is reflexive and transitive).*

In our approach, we intend to use more general relations than equivalence relations and different from preorder relations. Below, we recall a non-transitive indiscernibility relation that will be considered in this work. But first, we need to recall the notion of information system.

Definition 6. *An information system $(U, \mathcal{A})$ is a tuple, such that $U = \{x_1, x_2, \ldots, x_n\}$ and $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$ are finite, non-empty sets of objects and attributes, respectively. Each $a \in \mathcal{A}$ is associated with a mapping $\bar{a} \colon U \to V_a$, where V_a is the value set of a over U.*

If $V_a = \{0, 1\}$ for each $a \in \mathcal{A}$, we say that $(U, \mathcal{A})$ is a Boolean *information system.*

Now, we introduce the notion of s-indiscernibility relation.

Definition 7. *Given an information system $(U, \mathcal{A})$, $s \in \mathbb{N}$ and $B \subseteq \mathcal{A}$, the s-indiscernibility relation with respect to B, R_B^s, is defined as follows.*

Two objects $x, y \in U$ belongs to R_B^s if and only if there are at most s attributes $\{a_1, \dots, a_s\} \subseteq B$ such that $\bar{a}_k(x) \neq \bar{a}_k(y)$ for all $k \in \{1, \dots, s\}$.

If $(x, y) \in R_B^s$, we say that x and y are s-indiscernible in B. When $B = \mathcal{A}$, we simply say that x and y are s-indiscernible and the relation is denoted as R^s.

In this paper, we focus on the study of a special kind of information system called decision system.

Definition 8. *A decision system $(U, \mathcal{A} \cup \{d\})$ is a kind of information system in which $d \notin \mathcal{A}$ is called the decision attribute.*

In this framework, the notions of positive region and the degree of dependency are generalized as follows.

Definition 9. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system, $B \subseteq \mathcal{A}$ and (U, R_B) a derived approximation space. The R_B-left positive and R_B-right positive regions with respect to R_B, denoted as $POS_{R_B}^\ell$ and $POS_{R_B}^r$ respectively, are defined as:*

$$POS_{R_B}^\ell = \bigcup_{x \in U} R_B \uparrow_r \left(R_B \downarrow^\ell [x]_d \right)$$

$$POS_{R_B}^r = \bigcup_{x \in U} R_B \uparrow_\ell \left(R_B \downarrow^r [x]_d \right)$$

and the degree of dependency of d over R_B, $\gamma_{R_B}^*$, as:

$$\gamma_{R_B}^* = \frac{\max \left\{ \mathrm{Card}(POS_{R_B}^\ell), \mathrm{Card}(POS_{R_B}^r) \right\}}{\mathrm{Card}(U)}$$

where $[x]_d$ represents the equivalence class of the object $x \in U$ with respect to the indiscernibility relation Ind_d given by

$$\mathrm{Ind}_d = \{(x, y) \in U \times U \mid \bar{d}(x) = \bar{d}(y)\}$$

Remark 1. The degree of dependency $\gamma_{R_B}^* = 1$ plays a remarkable role in decision systems since in such a case, a perfect classification can be performed taking into account the information provided by the approximation space (U, R_B). Additionally, note that if $\gamma_{R_B}^* = 1$, we have

$$\max \left\{ \mathrm{Card}(POS_{R_B}^\ell), \mathrm{Card}(POS_{R_B}^r) \right\} = \mathrm{Card}(U).$$

In other words, $POS_{R_B}^\ell = U$ or $POS_{R_B}^r = U$. Moreover, by Theorem 1 we have that $R_B \uparrow_r \left(R_B \downarrow^\ell [x]_d \right) \subseteq [x]_d$ and $R_B \uparrow_\ell \left(R_B \downarrow^r [x]_d \right) \subseteq [x]_d$, as a result, if $POS_{R_B}^\ell = U$ we have that $R_B \uparrow_r \left(R_B \downarrow^\ell [x]_d \right) = [x]_d$, for all $x \in U$, and when $POS_{R_B}^r = U$ then $R_B \uparrow_\ell \left(R_B \downarrow^r [x]_d \right) = [x]_d$, for all $x \in U$.

3   Representative Set of Objects of a Decision Class

In this section, we introduce a novel class of objects, called representative. The underlying idea in such a definition is to determine a subset of objects that characterizes a certain decision class.

Definition 10. *Given a decision system* $(U, \mathcal{A} \cup \{d\})$ *and* $x \in U$, *we say that a subset of objects* $X \subseteq U$ *is:*

- *a left-representative set of the decision class* $[x]_d$ *if* $R \uparrow_\ell (X) = [x]_d$.
- *a right-representative set of the decision class* $[x]_d$ *if* $R \uparrow_r (X) = [x]_d$.

We will denote as $ROS^\ell([x]_d)$ *and* $ROS^r([x]_d)$ *to the set of left-representative sets and right-representative sets of the decision class* $[x]_d$, *respectively.*

Notice that when the relation R is symmetric, the left-representative sets coincide with the right-representative sets. In such a case, we call that sets *representative sets of a decision class* $[x]_d$, and denote the set formed by them as $ROS([x]_d)$.

Note also that if X is a representative set of a decision class $[x]_d$ for certain $x \in U$, then every element in the class of $[x]_d$ is related at least with one element in X and moreover, all the elements in X are related only to elements of $[x]_d$. In other words, we can characterize the elements in the class $[x]_d$ by checking which objects in U are related (or not) to elements in X.

The following example illustrates the previous definition.

Example 1. Consider a decision system $(U, \mathcal{A} \cup \{d\})$ composed of the set of objects $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, the set of attributes $\mathcal{A} = \{a_1, a_2, a_3, a_4, a_5\}$ related between them as the following table shows:

	a_1	a_2	a_3	a_4	a_5	d
x_1	x		x	x		x
x_2		x	x		x	
x_3	x	x	x	x		x
x_4	x	x	x		x	
x_5		x	x	x	x	
x_6	x	x	x			x
x_7		x	x		x	

Note that, in this case, the obtained decision classes are:

$$[x_1]_d = \{x_1, x_3, x_6\}$$
$$[x_2]_d = \{x_2, x_4, x_5\}$$

We consider $B = \mathcal{A}$ and the s-indiscernibility relation with $s = 1$, that is, $R_\mathcal{A}^1$. The results obtained from the considered s-indiscernibility relation is shown in the table below.

$R^1_{\mathcal{A}}$	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	X		X				
x_2		X		X	X		X
x_3	X		X			X	
x_4		X		X		X	X
x_5		X			X		X
x_6			X	X		X	
x_7		X		X	X		X

According to the previous table, we obtain that:

$$R \uparrow (\{x_3\}) = \{x_1, x_3, x_6\} = [x_1]_d$$
$$R \uparrow (\{x_2\}) = \{x_2, x_4, x_5\} = [x_2]_d$$

Therefore, we can assert that the set $\{x_3\}$ is a representative set of the decision class $[x_1]_d$ and the set $\{x_2\}$ is a representative set of the decision class $[x_2]_d$. But $\{x_3\}$ is not the only representative set of the decision class $[x_1]_d$, we can also find a set, composed of more than one object, that represents the same decision class as, for example, the set $\{x_1, x_3\}$; it is easy to check that $R \uparrow (\{x_1, x_3\}) = [x_1]_d$. However, it is important to note that the set $\{x_1\}$ is not a representative set of it decision class $[x_1]_d$ because $x_6 \notin R \uparrow (\{x_1\})$. In addition, the set $\{x_3, x_6\}$ is not a representative set of the decision class $[x_1]_d$ either, since $x_4 \in R \uparrow (\{x_3, x_6\})$ and $x_4 \notin [x_1]_d$. $\qquad\square$

In the previous example, we have shown that adding or removing objects from a given representative set of a decision class may change such a feature. Therefore, it looks interesting to study the structure of the set composed of all representative sets of a certain decision class.

In order to present the first result related to the structure of the representative sets, we need to introduce the following definition.

Definition 11. *Given an approximation space (U, R), we say that*

- *$x \in U$ is a left-isolated object of the relation R if there is no element $y \in U$ satisfying that $(x, y) \in R$. The set composed of all the left-isolated objects is denoted as $\mathrm{Is}^\ell(R)$*
- *$y \in U$ is a right-isolated object of the relation R if there is no element $x \in U$ satisfying that $(x, y) \in R$. The set composed of all the right-isolated objects is denoted as $\mathrm{Is}^r(R)$.*
- *$x \in U$ is a isolated object of the relation R if it is left-isolated and right-isolated. The set composed of all the isolated objects is denoted as $\mathrm{Is}(R)$.*

The first result shows that the set of left-representative (right-representative) sets of two different decision classes are disjoint except for isolated objects.

Proposition 1. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system, R a discernibility relation and $X, Y \subseteq U$ two left-representative (right-representative) sets of two different decision classes, then the set $X \bigcap Y$ is contained in the set of left-isolated (right-isolated) objects of the relation R.*

Proof. Let $x, y \in U$ such that $[x]_d \neq [y]_d$. Then, necessarily $[x]_d \bigcap [y]_d = \varnothing$. Let X and Y be left-representative sets of the classes $[x]_d$ and $[y]_d$, respectively, and let us prove that $X \bigcap Y \subseteq \mathrm{Is}^\ell(R)$. We consider $x \in X \bigcap Y$, then we have that $R \!\uparrow_\ell (x) \subseteq R \!\uparrow_\ell (X \bigcap Y)$. In addition, since $(R \!\uparrow_\ell, R \!\downarrow^r)$ is a Galois connection, we have that

$$R \!\uparrow_\ell (x) \subseteq R \!\uparrow_\ell \left(X \bigcap Y \right) \subseteq R \!\uparrow_\ell (X) \bigcap R \!\uparrow_\ell (Y) = [x]_d \bigcap [y]_d = \varnothing.$$

Therefore, we have that $R \!\uparrow_\ell (x) = \varnothing$. Then, according to Definition 3, there is not $y \in U$ such that $(x, y) \in R$, that is, x is a left-isolated object of the relation R. Hence, we have that $X \bigcap Y \subseteq \mathrm{Is}^\ell(R)$.

The proof with right-representative sets is developed in an analogous way. $\quad\square$

The following result shows that the set of representative sets of a certain decision class has the structure of a join-semilattice with respect to the standard ordering between subsets.

Proposition 2. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system and let $X, Y \subseteq U$ such that X is a left-representative (right-representative) set of a decision class $[x]_d$, with $x \in U$, and $R \!\uparrow_\ell (Y) \subseteq [x]_d$ (respectively $R \!\uparrow_r (Y) \subseteq [x]_d$). Then $X \bigcup Y$ is also a left-representative (right-representative) set of $[x]_d$.*

Proof. Let $X, Y \subseteq U$ such that X is a left-representative set of a decision class $[x]_d$, with $x \in U$, and $R \!\uparrow_\ell (Y) \subseteq [x]_d$. Since $(R \!\uparrow_\ell, R \!\downarrow^r)$ is a Galois connection, we have that

$$R \!\uparrow_\ell \left(X \bigcup Y \right) = R \!\uparrow_\ell (X) \bigcup R \!\uparrow_\ell (Y) = [x]_d \bigcup R \!\uparrow_\ell (Y) = [x]_d.$$

In other words, $X \bigcup Y$ is a left-representative set of $[x]_d$.

The proof follows similarly for the right-representative sets. $\quad\square$

The following consequence of the previous proposition shows that the construction of left-representative sets and right-representative sets can be done by singletons.

Corollary 1. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system and let $X, Y \subseteq U$ such that X and $X \bigcup Y$ are left-representative (right-representative) sets of a decision class $[x]_d$, with $x \in U$. Then $X \bigcup \{y\}$ is also a left-representative (right-representative) set of $[x]_d$, for all $y \in Y$.*

The following result shows that the set of representative sets of a certain decision class has the structure of join-semilattice with respect to the standard ordering between subsets.

Corollary 2. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system and let $X, Y \subseteq U$ be two left-representative (right-representative) sets of the same decision class $[x]_d$, with $x \in U$. Then $X \bigcup Y$ is also a left-representative (right-representative) set of $[x]_d$.*

Example 2. In Example 1, we have two representative sets for the class $[x_1]_d$:

$$ROS([x_1]_d) = \big\{\{x_3\}, \{x_1, x_3\}\big\}.$$

On the other hand, we have six representative sets for the class $[x_2]_d$, namely:

$$ROS([x_2]_d) = \big\{\{x_2\}, \{x_7\}, \{x_2, x_7\}, \{x_2, x_5\}, \{x_7, x_5\}, \{x_2, x_5, x_7\}\big\}.$$

Note that $ROS([x_1]_d)$ and $ROS([x_2]_d)$ are disjoint, because there is not isolated elements in the considered discernibility relation, as Proposition 1 asserts.

On the other hand, according to Proposition 2, it is easy to check the join of arbitrary representative sets is also a representative for the respective class. Specifically, we can observe that the representative sets for the class $[x_1]_d$ has a lattice structure, but the set of representative sets for the class $[x_2]_d$ only has the structure of a join-semilattice. That fact can be seen, for example, in the intersection of the sets $\{x_2, x_5\}$ and $\{x_7, x_5\}$ which is the singleton $\{x_5\}$ that is not a representative set of the class $[x_2]_d$.   □

Let us analyze now the minimal and maximal representative sets of decision classes.

Definition 12. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system and $X \subseteq U$ a representative set of a decision class $[x]_d$, with $x \in U$. We say that:*

- *X is a minimal left-representative (right-representative) set of the decision class $[x]_d$ if $X \setminus \{x'\}$ is not a left-representative (right-representative) set of $[x]_d$, for all $x' \in X$.*
- *X is a maximal left-representative (right-representative) set of the decision class $[x]_d$ if there is no object $x' \in U \setminus X$ such that $X \bigcup \{x'\}$ is a left-representative (right-representative) set of $[x]_d$.*

Thanks to the join-semilattice structure of the set of representative sets of a decision class (Proposition 2), we can directly infer that the maximal representative set of a decision class is unique; i.e., it is a maximum, as the following corollary states.

Corollary 3. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system and $x \in U$. If there exists a left-representative (right-representative) set of a decision class $[x]_d$ then, there is a unique maximal left-representative (right-representative) set of that decision class.*

The unicity stated by the previous result does not hold for minimal representative sets; i.e., it may exists several minimal representative sets of a decision class. The following example shows that fact.

Example 3. Coming back to Example 2, according to Definition 12, we have that the sets $\{x_3\}$ is the only minimal representative set of the decision class $[x_1]_d$. However, there are two minimal representative sets of the decision class $[x_2]_d$, namely $\{x_2\}$ and $\{x_7\}$.

On the other hand, it can be proved easily that $\{x_1, x_3\}$ and $\{x_2, x_5, x_7\}$ are the two maximal representative sets of $[x_1]_d$ and $[x_2]_d$, respectively.   □

The following result determines the maximal left-representative set of a decision class, if it exists.

Theorem 3. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system, $B \subseteq \mathcal{A}$, (U, R_B) an approximation space and $x \in U$. Then:*

– *If there exists a right-representative set of the decision class $[x]_d$, then the set $R \downarrow^\ell ([x]_d)$ is the maximum right-representative set of the decision class $[x]_d$.*
– *If there exists a left-representative set of the decision class $[x]_d$, then the set $R \downarrow^r ([x]_d)$ is the maximum left-representative set of the decision class $[x]_d$.*

Proof. Let $x \in U$ such that there exists a right-representative set of its decision class. By Corollary 3, we have that there exists the maximum right-representative set of the decision class $[x]_d$, denoted by $X \subseteq U$. By Theorem 1, we have that $R_B \uparrow_r \left(R_B \downarrow^\ell [x]_d \right) \subseteq [x]_d$. Then, since X is a right-representative set of the decision class $[x]_d$, by Proposition 2, we have that $X \cup R_B \downarrow^\ell [x]_d$ is a representative set of the decision class $[x]_d$ as well. As a result, by the maximality of the set X, we have that $R_B \downarrow^\ell [x]_d \subseteq X$.

Let us prove now that $X \subseteq R_B \downarrow^\ell [x]_d$. Consider $y \in X$, since X is a right-representative set of $[x]_d$ and by the monotonicity of $R \uparrow_r$, we have that $R \uparrow_r (\{y\}) \subseteq [x]_d$. Therefore, by definition of $R \uparrow_r$, if we consider $z \in U$ satisfying that $zR \cap \{y\} \neq \varnothing$, then $z \in [x]_d$. As a consequence, for all $z \in U$ such that $(z, y) \in R$ we have that $z \in [x]_d$, which is equivalent to say that $Ry \subseteq [x]_d$ and then, $y \in R_B \downarrow^\ell [x]_d$. In other words $X \subseteq R_B \downarrow^\ell [x]_d$.

Finally, we can assert that $R_B \downarrow^\ell [x]_d$ is the maximal right-representative set of $[x]_d$.

The proof for the maximal left-representative set of $[x]_d$ follows analogously. □

In the last result we relate the representative sets of a decision class to the degree of dependency $\gamma^*_{R_B}$. Specifically, we show that $\gamma^*_{R_B} = 1$ is equivalent to assert the existence of left-representative sets for each decision class or right-representative sets for each decision class.

Theorem 4. *Let $(U, \mathcal{A} \cup \{d\})$ be a decision system, $B \subseteq \mathcal{A}$ and (U, R_B) an approximation space. R_B satisfies that $\gamma^*_{R_B} = 1$ if and only if*

– *there exists at least one left-representative set for each decision class,*
– *or there exists at least one right-representative set for each decision class.*

Proof. Let us assume that R_B satisfies that $\gamma^*_{R_B} = 1$. Then, $R_B \uparrow_r \left(R_B \downarrow^\ell [x]_d \right) = [x]_d$, for all $x \in U$ or $R_B \uparrow_\ell \left(R_B \downarrow^r [x]_d \right) = [x]_d$ for all $x \in U$ (see Remark 1). Therefore, we have that $R_B \downarrow^\ell [x]_d$ is a right-representative set of $[x]_d$, for all $x \in U$, or $R_B \downarrow^r [x]_d$ is a left-representative set of $[x]_d$, for all $x \in U$.

Now, let us prove the converse. Without loss of generality, let us assume that there exists at least one left-representative set for each decision class. Then, by Theorem 3, we have that $R \downarrow^r ([x]_d)$ is the maximal left-representative set of the class $[x]_d$, that is, $R_B \uparrow_\ell (R_B \downarrow^r [x]_d) = [x]_d$ for all $x \in U$. As a consequence:

$$POS^r_{R_B} = \bigcup_{x \in U} R_B \uparrow_\ell (R_B \downarrow^r [x]_d) = \bigcup_{x \in U} [x]_d = U$$

and therefore, $\gamma^*_{R_B} = 1$.   □

4   Conclusions and Future Work

In this paper we have provided the formal definition of the notion of representative set of objects of a decision class. Moreover, we have presented some mathematical properties of such kind of sets and shown its connection with a classification task based on rough sets.

There are different future lines based on the notion of representative set of objects. Firstly, the obtention of more mathematical properties about the objects forming representative sets is interesting for several purposes; for example, for its construction or for determining minimal representative sets. Secondly, analyzing the relationship between reducts of attributes and the set of representative sets of objects has our attention as well. Last but not least, the construction of a classification procedure based on representative sets of objects seems to be appropriated when the dataset is involved with uncertainty; for example when we need to classify an object that is discernible with all the objects in the training dataset.

References

1. Benítez-Caballero, M.J., Medina, J., Ramírez-Poussa, E., Ślęzak, D.: Bireducts with tolerance relations. Inf. Sci. **435**, 26–39 (2018)
2. Benítez-Caballero, M.J., Medina, J., Ramírez-Poussa, E., Ślęzak,D.: Rough-set-driven approach for attribute reduction in fuzzy formal concept analysis. Fuzzy Sets Syst. (2019)
3. Birkhoff, G.: Lattice Theory, 3rd edn. American Mathematical Society, Providence (1967)
4. Bloch, I.: On links between mathematical morphology and rough sets. Pattern Recogn. **33**, 1487–1496 (2000)
5. Cornelis, C., Jensen, R., Hurtado, G., Ślęzak, D.: Attribute selection with fuzzy decision reducts. Inf. Sci. **180**, 209–224 (2010)
6. Cornelis, C., Medina, J., Verbiest, N.: Multi-adjoint fuzzy rough sets: definition, properties and attribute selection. Int. J. Approx. Reason. **55**, 412–426 (2014)

7. Couso, I., Dubois, D.: Rough sets, coverings and incomplete information. Fundam. Inf. **108**(3–4), 223–247 (2011)
8. Davey, B., Priestley, H.: Introduction to Lattices and Order, 2nd edn. Cambridge University Press, Cambridge (2002)
9. Denecke, K., Erné, M., Wismath, S.L. (eds.): Galois Connections and Applications. Kluwer Academic Publishers, Dordrecht (2004)
10. Han, S.-E.: Roughness measures of locally finite covering rough sets. Int. J. Approx. Reason. **105**, 368–385 (2019)
11. Hassanien, A.E., Schaefer, G., Darwish, A.: Computational intelligence in speech and audio processing: recent advances. In: Gao, X.-Z., Gaspar-Cunha, A., Köppen, M., Schaefer, G., Wang, J. (eds.) Soft Computing in Industrial Applications. Advances in Intelligent and Soft Computing, vol. 75, pp. 303–311. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-11282-9_32
12. Janusz, A., Ślęzak, D.: Rough set methods for attribute clustering and selection. Appl. Artif. Intell. **28**(3), 220–242 (2014)
13. Janusz, A., Ślęzak, D., Nguyen, H.S.: Unsupervised similarity learning from textual data. Fundam. Inf. **119**, 319–336 (2012)
14. Järvinen, J., Radeleczki, S., Veres, L.: Rough sets determined by quasiorders. Order **26**(4), 337–355 (2009)
15. Kudo,Y., Murai, T.: An attempt of object reduction in rough set theory. In: 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS), pp. 33–36, December 2018
16. Mac Parthalain, N., Jensen, R.: Simultaneous feature and instance selection using fuzzy-rough bireducts. In: 2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2013), pp. 1–8, July 2013
17. Madrid, N., Medina, J., Ramírez-Poussa, E.: Rough sets based on galois connections. Int. J. Appl. Math. (2020, accepted)
18. Medina, J.: Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. Comput. Math. Appl. **64**(6), 1992–2002 (2012)
19. Pagliani, P.: The relational construction of conceptual patterns - tools, implementation and theory. In: Kryszkiewicz, M., Cornelis, C., Ciucci, D., Medina-Moreno, J., Motoda, H., Raś, Z.W. (eds.) Rough Sets and Intelligent Systems Paradigms. LNCS, vol. 8537, pp. 14–27. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08729-0_2
20. Pagliani, P.: Covering rough sets and formal topology - a uniform approach through intensional and extensional constructors. In: Peters, J., Skowron, A. (eds.) Transactions on Rough Sets XX. LNCS, vol. 10020, pp. 109–145. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53611-7_4
21. Shao, M.-W., Liu, M., Zhang, W.-X.: Set approximations in fuzzy formal concept analysis. Fuzzy Sets Syst. **158**(23), 2627–2640 (2007). Theme: Logic
22. Stawicki, S., Ślęzak, D.: Recent advances in decision bireducts: complexity, heuristics and streams. In: Lingras, P., Wolski, M., Cornelis, C., Mitra, S., Wasilewski, P. (eds.) RSKT 2013. LNCS, vol. 8171, pp. 200–212. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41299-8_19
23. Stawicki, S., Ślęzak, D., Janusz, A., Widz, S.: Decision bireducts and decision reducts - a comparison. Int. J. Approx. Reason. **84**, 75–109 (2017)
24. Varma, P.R.K., Kumari, V.V., Kumar, S.S.: A novel rough set attribute reduction based on ant colony optimisation. Int. J. Intell. Syst. Technol. Appl. **14**(3–4), 330–353 (2015)

25. Yao, Y., Lingras, P.: Interpretations of belief functions in the theory of rough sets. Inf. Sci. **104**(1), 81–106 (1998)
26. Yao, Y., Yao, B.: Covering based rough set approximations. Inf. Sci. **200**, 91–107 (2012)
27. Yao, Y.Y.: Two views of the theory of rough sets in finite universes. Int. J. Approx. Reason. **15**(4), 291–317 (1996). Rough Sets
28. Ziarko, W.: Probabilistic approach to rough sets. Int. J. Approx. Reason. **49**(2), 272–284 (2008). Special Section on Probabilistic Rough Sets and Special Section on PGM'06