δ-information reducts and bireducts[*]

Mª José Benítez,[1] Jesús Medina,[1] Dominik Ślęzak[2]

[1]Department of Mathematics, University of Cádiz. Spain
Email: `jesus.medina@uca.es`
[2]Institute of Mathematics, University of Warsaw. Poland
Email: `slezak@mimuw.edu.pl`

Abstract

Attribute reduction is an important step in order to decrease the computational complexity to derive information from databases. In this paper, we extend the notions of reducts and bireducts introduced in rough set theory for attribute reduction purposes and let them work with similarity relations defined on attribute values. Hence, the related mathematical concepts will be introduced and the characterizations of the new reducts and bireducts will be given in terms of the corresponding generalizations of the discernibility function.

Keywords: Similarity relation; information system; reduct; bireduct; discernibility function.

1. Introduction

Rough set theory was proposed by Pawlak [16] as a tool for modelling and processing incomplete information in information systems. In principle, it uses subsets of attributes to approximate subsets of objects. However, one can also extend it to achieve deeper duality between objects and attributes [4]. This kind of generalization of rough set theory refers also to property-oriented concept lattices discussed in [2].

Another important direction was to integrate fuzzy sets. A first definition, the rough fuzzy sets, was given by Fariñas del Cerro and Prade in the eighties [5], and after that one, numerous hybrid models have been introduced [13, 14, 15, 17]. Such mathematical techniques are important in the extraction and manipulation of information in relational databases with inaccurate, missing or lost data.

Attribute reduction in such frameworks is a step in decreasing the computational complexity of knowledge discovery. The main idea is to reduce the size of the database, without losing information about analyzed elements. To this end the reducts were studied in a number of papers, e.g. [3, 9, 10, 12, 23, 25].

However, this kind of reduction is not enough sometimes. In order to provide more flexibility, we can further reduce attributes so they do not keep the complete information about objects, and, in parallel, we can register objects for which information is lost. This is a new kind of duality between subsets of attributes and subsets of objects in the attribute reduction process. The pairs of these searched subsets are called bireducts and have recently become important in rough set theory [7, 11, 20, 21].

Computing the bireducts is more complex but it also provides more powerful tools for representation of data dependencies. In order to further extend expressive power of bireducts, in this paper, we study them within an environment where the notion of equality is replaced by the notion of similarity between attribute values. This similarity, in the case of numeric database, can be obtained from a classic distance.

Similarity relations give us a comparison between attributes, letting create a gradual hierarchy of values for each attribute [8, 19, 22]. These relationships allow us to adapt our study to different situations depending on our needs.

Consequently, we introduce the notions of δ-information reduct and bireduct in a similarity environment where the corresponding reducts and bireducts are characterized by cubes of discernibility functions. This may be useful in cases with complex information system. Complex attribute values occur often in the medical domain like data representing medical treatment of patiens with the head and neck cancer cases presented in [1]. Complex attribute values can also appear in other medical studies, as well as, multimedia or robotics problems.

The organization of the paper is the following. Some basic concepts related to classical theory of propositional logic and the notion of similarity relation are recalled in Section 2. Section 3 presents the basic definitions in rough set theory, the notions of δ-similar and δ-discordant, the corresponding definition of δ-information reduct and its characterization by the extended discernibility function. The bireducts in the new similarity environment are introduced in Section 4. Conclusions and prospects for future work are given in Section 5.

2. Preliminaries

In this paper the classical theory of propositional logic will be considered in order to interpret the expression of the discernibility function. Hence, sev-

---

eral basic notions of propositional logic will be recalled.

An alphabet $\mathcal{A}$ is formed by a numerable set of symbols or propositional variables

$$\mathcal{V} = \{p_1, q_1, r_1 \ldots, p_2, q_2, r_2, \ldots, p_n, q_n, r_n, \ldots\}$$

the constant symbols $\top$ and $\bot$, the symbols $\neg, \wedge, \vee, \rightarrow$ and $\leftrightarrow$, called *connectives* or *logical operators*, and the punction symbols "(", ")".

The language $\mathcal{L}$ is the free generated inductive closing of the base set $\mathcal{V} \cup \{\top, \bot\}$ for the constructors $C_\neg, C_\wedge, C_\vee, C_\rightarrow$ and $C_\leftrightarrow$ defined as follows: For whichever two chains X and Y of the alphabet $\mathcal{A}$, we define:

$$\begin{aligned}
C_\neg(X) &= \neg X \\
C_\wedge(X, Y) &= (X \wedge Y) \\
C_\vee(X, Y) &= (X \vee Y) \\
C_\rightarrow(X) &= (X \rightarrow Y) \\
C_\leftrightarrow(X) &= (X \leftrightarrow Y)
\end{aligned}$$

$\top$ is read as "true" and $\bot$ is read as "false" and, if A and B are well-formed formulas (WFF),

- $\neg A$ is read as "no A" and is called *the negation of A*.
- $A \wedge B$ is read as "A and B" and is called *conjunction* of A and B.
- $A \vee B$ is read as "A or B" and is called *disjunction* of A and B.
- $A \rightarrow B$ is read as "If A then B" and is called *implication* with antecedent A and consequent B.
- $A \leftrightarrow B$ is read as "A if and only if B" and is called *biimplication* of A and B.

Therefore the *language of the propositional logic* is formed by the alphabet and the set of WFFs.

Definition 1 *The propositional symbols, with their negations are called* literals. *We say that the literals p and $\neg p$ are* complementary literals.

If l is a literal, we are going to denote its complementary as $\bar{l}$.

Definition 2 *A WFF is a* cube *if it is $\top$, $\bot$, a literal or a conjunction (possibly empty) of literals. We say that it is a* restricted cube *if it does not contain repeated literals neither pairs of complementary literals.*

Definition 3 *A WFF is a* clause *if it is $\top$, $\bot$, a literal or a disjunction (possibly empty) of literals. We say that it is a* restricted clause *if it does not contain repeated literals neither pairs of complementary literals.*

Definition 4 *A WFF is said to be in* disjunctive normal form *(DNF) if it is: $\top$, $\bot$, a cube or a disjunction (possibly empty) of cubes.*

A WFF is said to be in conjunctive normal form *(CNF) if it is: $\top$, $\bot$, a clause or conjunction (possibly empty) of clause.*

Disjunctive and conjunctive normal forms may be reduced using absorption laws until none of them can be further reduced, obtaining the reduced forms:

Definition 5 *A DNF is said to be* restricted *(briefly, RDNF), if it satisfies that any cube contains a literal or its complementary, and it does not contain repeated literals and other cubes.*

A CNF is said to be restricted *(briefly, RCNF), if it satisfies that any clause contains a literal or its complementary, and it does not contain repeated literals and other clauses.*

Example 6 *The following example illustrate the the previous concepts:*

- *The formula $(a \wedge b) \vee (d \wedge e) \vee e$ is a DNF, but it is not a RDNF, since the absorption law can be applied and a the new formula is obtained: $(a \wedge b) \vee e$, which is just a RDNF.*

- *The CNF $(a \vee b) \wedge a \wedge (b \vee d)$ is not restricted (RCNF). Applying the absorption law we obtain the equivalent formula: $a \wedge (b \vee d)$, which already is a RCNF.*

The notion of RDNF is important in order to introduce and manage discernibility functions used in rough set theory. It will be also used to generalize our framework to consider similarity relations.

Hence, we will continue recalling the definition of similarity relation, which extends the notion of equivalence relation and therefore the concept of equality.

Definition 7 *Given an arbitrary set V, the mapping $E_V \colon V \times V \to [0,1]$, is called* fuzzy similarity relation *if the following properties hold:*

1. *E is reflexive.*
2. *E is symmetric.*
3. *E is transitive, that is, $E(v_1, v_2) \wedge E(v_2, v_3) \leq E(v_1, v_2)$, for all $v_1, v_2, v_3 \in V$.*

In theory, one can define a similarity relation over the set of objects in an arbitrary way. However, in practice it is indeed reasonable to refer to values of objects for available attributes.

3. Generalization of reducts by similarities: δ-information reducts

First of all, we will recall the basic definitions of Rough Set Theory and the use of a similarity relation will provide the notions of δ-similar and δ-discordant objects, with respect to a threshold δ.

Definition 8 *An information system $(U, \mathcal{A})$, where $U = \{1, \ldots, n\}$ and $\mathcal{A} = \{a_1, \ldots, a_m\}$ are finite, non-empty sets of objects and attributes, respectively. Each a in $\mathcal{A}$ corresponds to a mapping $\bar{a} \colon U \to V_a$, where V_a is the value set of a over U.*

Next, an example of information system is introduced.

Example 9 *Let be the information system* $\mathbb{A} = (U, \mathcal{A})$, *where the set of objects is* $U = \{1, 2, 3, 4, 5, 6\}$, *the set of attributes is* $A = \{Outlook, Temp., Humid., Wind\}$ *and the following table shows the relationship between them:*

	Outlook	Temp.	Humid.	Wind
1	sunny	hot	high	weak
2	sunny	hot	high	strong
3	overcast	hot	high	weak
4	rain	mild	high	weak
5	rain	cool	normal	weak
6	rain	cool	normal	strong

Thus, we obtain four functions associated with each of the attributes whose domains are:

$V_O = \{sunny,\ overcast,\ rain\}$ $\quad V_T = \{hot,\ mild,\ cool\}$
$V_H = \{high,\ normal\}$ $\qquad\qquad V_W = \{weak,\ strong\}$

An important relationship among the object is needed.

Definition 10 *For every subset B of $\mathcal{A}$, the B-indiscernibility relation[1] I_B is defined as the equivalence relation given by the following set of pairs:*

$$\{(i,j) \in U \times U \mid \text{ for all } a \in B,\ \bar{a}(i) = \bar{a}(j)\}$$

Then, I_B is an equivalence relation, where each class can be written as $[i]_B = \{j \mid (i,j) \in I_B\}$ and produces a partition on U denoted as: $U/I_B = \{[i]_B \mid i \in U\}$.

In rough set theory, data is represented as an information system. Given $X \subseteq U$, its lower and upper approximation w.r.t. B are defined by

$$I_B{\downarrow}X = \{i \in U \mid [i]_B \subseteq X\} \quad (1)$$
$$I_B{\uparrow}X = \{i \in U \mid [i]_B \cap X \neq \varnothing\} \quad (2)$$

The following example shows the classes of the previous particular information system.

Example 11 *From the information system in Example 9, the $\mathcal{A}$-indiscernibility relation is defined in the following table.*

$I_{\mathcal{A}}$	1	2	3	4	5	6
1	1	0	0	0	0	0
2	0	1	0	0	0	0
3	0	0	1	0	0	0
4	0	0	0	1	0	0
5	0	0	0	0	1	0
6	0	0	0	0	0	1

Therefore, trivial equivalence classes are obtained: $[i]_{\mathcal{A}} = \{i\}$, for all $i \in U$. Hence, in this case, $I_{\mathcal{A}}{\downarrow}X = I_{\mathcal{A}}{\uparrow}X$ for all $X \subseteq U$.

Another example is given from the information system $(X, \mathcal{A})$ where X, $\mathcal{A}$ and the mappings

	Temperature	Headache
x_1	Hight	Yes
x_2	Normal	Yes
x_3	Hight	Yes
x_4	Normal	No

Table 1: Data of the information system

$\bar{a} \colon X \to V_a$, *where V_a is the value set of a over X, for all $a \in \mathcal{A}$ are shown in Table 11. The $\mathcal{A}$-indiscernibility relation of this information system is given in Table 2. From this equivalence relation,*

$I_{\mathcal{A}}$	x_1	x_2	x_3	x_4
x_1	1	0	1	0
x_2	0	1	0	0
x_3	1	0	1	0
x_4	0	0	0	1

Table 2: $\mathcal{A}$-indiscernibility relation

the non-trivial obtained classes are:

- $[x_1]_{\mathcal{A}} = \{x_1, x_3\} = [x_3]_{\mathcal{A}}$
- $[x_2]_{\mathcal{A}} = \{x_2\}$
- $[x_4]_{\mathcal{A}} = \{x_4\}$

Given $A_1 = \{x_2, x_3\}$, we have that

$$I_{\mathcal{A}}{\downarrow}A_1 = \{x_2\}$$
$$I_{\mathcal{A}}{\uparrow}A_1 = \{x_1, x_2, x_3\}$$

If $A_2 = \{x_1, x_3\}$, then

$$I_{\mathcal{A}}{\downarrow}A_1 = \{x_1, x_3\}$$
$$I_{\mathcal{A}}{\uparrow}A_1 = \{x_1, x_3\}$$

The notion of reduct is fundamental in this paper.

Definition 12 *Given an information system $\mathbb{A} = (U, \mathcal{A})$, the set $B \subseteq A$ is called* information reduct *if and only it satisfies $I_B = I_{\mathcal{A}}$ and $I_{B \setminus \{a\}} \neq I_{\mathcal{A}}$, for all $a \in B$.*

A well-known approach to generate all reducts of an information system is based on its discernibility matrix and function [18]. The discernibility matrix of $(U, \mathcal{A})$ is the $n \times n$ matrix O, defined by, for i and j in $\{1, ..., n\}$,

$$O_{ij} = \{a \in \mathcal{A} \mid \bar{a}(i) \neq \bar{a}(j)\} \quad (3)$$

The *discernibility function* of $(U, \mathcal{A})$ is the map $f \colon \{0,1\}^m \to \{0,1\}$, defined by

$$f(a_1^*, ..., a_m^*) =$$
$$= \bigwedge \left\{ \bigvee O_{ij}^* \mid 1 \leq i < j \leq n \text{ and } O_{ij} \neq \varnothing \right\}$$

---

[1]When $B = \{a\}$, i.e., B is a singleton, we will write I_a instead of $I_{\{a\}}$.

in which $O_{ij}^* = \{a^* \mid a \in O_{ij}\}$. The Boolean variables $a_1^*, \ldots, a_m^*$ correspond to the attributes from $\mathcal{A}$. It can be shown that the prime implicants of f constitute exactly all decision reducts of $(U, \mathcal{A})$.

There are several possibilities to define a similarity relation on the set of objects U, $E_U \colon U \times U \to [0, 1]$. One of the most popular ways is from a family of similarity relations $\mathcal{E} = \{E_a \colon V_a \times V_a \to [0, 1] \mid a \in \mathcal{A}\}$ as follows:

$$E_U(i, j) = \bigwedge_{a \in A} (E_a(a(i), a(j))) \qquad (4)$$

Definition 13 *Given an information system $\mathbb{A} = (U, \mathcal{A})$, a pair $i, j \in U$ and a similarity relation family $\mathcal{E} = \{E_a \colon V_a \times V_a \to [0, 1] \mid a \in \mathcal{A}\}$, a pair of objects $i, j \in U$ is called δ-**similar** if for all $a \in \mathcal{A}$ we have:*

$$\delta \le E_a(a(i), a(j))$$

*with $\delta \in [0, 1]$. Otherwise, we say that objects $i, j \in U$ are δ-**discordant**, that is, the following holds:*

$$\{a \in \mathcal{A} \mid E_a(a(i), a(j)) < \delta\} \ne \varnothing$$

An equivalence relation can be defined associated with the notion of δ-similar.

Definition 14 *Given an information system $\mathbb{A} = (U, \mathcal{A})$, a similarity relation family $\mathcal{E} = \{E_a \colon V_a \times V_a \to [0, 1] \mid a \in \mathcal{A}\}$ and a threshold δ, a relation $S_{\mathcal{E}, \delta} = \{(i, j) \in U \times U \mid \delta \le E_a(a(i), a(j)), \text{ for all } a \in \mathcal{A}\}$ is defined.*

Lemma 15 *The relation $S_{\mathcal{E}, \delta}$ is an equivalence relation.*

As a consequence, a quotient set is defined on the set of objects U, which will be denoted as $\mathcal{U} = U/S_{\mathcal{E}, \delta}$. The equivalence classes of this quotient set will play an important role in the generalizations of the discernibility functions introduced in the next section.

In this section a threshold $\delta \in [0, 1]$ is fixed, from which we will use the notions of δ-similar and δ-discordant to define the generalization of the discernibility function using similarity relations.

Hence, an information system $\mathbb{A} = (U, \mathcal{A})$ and a similarity relation family $\mathcal{E} = \{E_a \colon V_a \times V_a \to [0, 1] \mid a \in \mathcal{A}\}$ will also be fixed. Moreover, a linear ordering $\le$ will also be fixed in U. Since the specific definition of the ordering is not important, any one can be considered. Given $i, j \in U$, we will say that $i < j$, if $i \le j$ and they are not the same object.

Below, the definitions of information reducts and bireducts based on a threshold $\delta \in [0, 1]$ are introduced.

Definition 16 *The set $B \subseteq \mathcal{A}$ is called δ-information reduct if and only if it is an irreducible subset such that every pair $i, j \in U$, which is δ-discordant by $\mathcal{A}$, is also δ-discordant by B.*

Definition 17 *The pair (B, X), where $B \subseteq \mathcal{A}$ and $X \subseteq U$, is called δ-information bireduct if and only if all pair i, j of X are δ-discordant by B and the following properties hold:*

1. *There is no $C \subsetneq B$ such that all pair $i, j \in X$ are δ-discordant by C.*
2. *There is no $X \subsetneq Y$ such that all pair $i, j \in Y$ are δ-discordant by B.*

Now, we are going to introduce the discernibility function in this general framework in order to obtain both δ-information reducts. Since for δ-information reducts only the attributes are needed we will call it unidimensional δ-discernibility function (uni δ-d function) and, in the next section, for δ-information bireducts, both attributes and objects are considered and so, we will call it bidimensional δ-discernibility function (bi δ-d function).

Definition 18 *Let $\mathbb{A} = (U, \mathcal{A})$ be a information system, the* unidimensional δ-discernibility function *of $\mathbb{A}$, is defined as the following conjunctive normal form (CNF):*

$$\tau_{\mathcal{A}}^{uni} = \bigwedge \left\{ \bigvee \{a \in \mathcal{A} \mid E_a(a(i), a(j)) < \delta\} \mid i, j \in U \right\}$$

where the elements of $\mathcal{A}$ are the propositional symbols of the language.

In order to compute this formula, a generalization of the discernibility matrix will be useful.

Definition 19 *Given an information system $\mathbb{A} = (U, \mathcal{A})$ and a threshold $\delta \in [0, 1]$, the symmetric matrix O_δ obtained from*

$$(O_\delta)_{ij} = \{a \in \mathcal{A} \mid E_a(a(i), a(j)) < \delta\} \qquad (5)$$

for i and j in $\{1, \ldots, n\}$, is called the δ-discernibility matrix associated with $\mathbb{A}$.

The following example shows a δ-discernibility matrix associated with the information system in Example 9 from a similarity relation.

Example 20 *First of all, we introduce a similarity relation for each attribute of the information system in Example 9 , which form the similarity relation family $\mathcal{E}$:*

$$\begin{array}{ll}
E_O : V_O \times V_O \to [0, 1] & E_T : V_T \times V_T \to [0, 1] \\
\quad E_O(s, o) = 0 & \quad E_T(h, m) = 0 \\
\quad E_O(s, r) = 0 & \quad E_T(h, c) = 0 \\
\quad E_O(o, r) = 0.5 & \quad E_T(m, c) = 0.5
\end{array}$$

$$\begin{array}{ll}
E_H : V_H \times V_H \to [0, 1] & E_W : V_W \times V_W \to [0, 1] \\
\quad E_H(h, n) = 0.25 & \quad E_W(w, s) = 0
\end{array}$$

The threshold of the δ-discernibility matrix will be fixed to $\delta = 0.3$. To build the 0.3-discernibility matrix, $O_{0.3}$ we must consider that it is a symmetric matrix and we can obtain the element $\varnothing$ for $O_\delta(i, j)$ in different ways:

a) *i and j are the same element;*
b) *There is no attribute satisfying the strict inequality $E_a(a(i), a(j)) < \delta$:*

$$
\begin{pmatrix}
\varnothing & & & & & \\
\{W\} & \varnothing & & & & \\
\{O\} & \{O, W\} & \varnothing & & & \\
\{O, T\} & \{O, T, W\} & \{T\} & \varnothing & & \\
\{O, T, H\} & \{O, T, H, W\} & \{T, H\} & \{H\} & \varnothing & \\
\{O, T, H, W\} & \{O, T, H\} & \{T, H, W\} & \{H, W\} & \{W\} & \varnothing
\end{pmatrix}
$$

Next, the characterization of the δ-information reducts is given.

Theorem 21 *Given a Boolean information system $\mathbb{A} = (U, \mathcal{A})$. An arbitrary set B, where $B \subseteq \mathcal{A}$, is a δ-information reduct of $\mathbb{A}$ if and only if the cube $\bigwedge_{b \in B} b$ is a cube in the RDNF of $\tau_{\mathcal{A}}^{uni}$.*

This characterization is considered in the following example.

Example 22 *From the information system of Example 22, the 0.3-discernibility matrix in Example 20, Definition 18 the unidimensional 0.3-discernibility function is:*

$$
\begin{aligned}
\tau^{uni} &= \{W\} \wedge \{O\} \wedge \{O \vee T\} \wedge \{O \vee T \vee H\} \\
&\wedge \{O \vee T \vee H \vee W\} \wedge \{O \vee W\} \\
&\wedge \{O \vee T \vee W\} \wedge \{T\} \wedge \{T \vee H\} \\
&\wedge \{T \vee H \vee W\} \wedge \{H\} \wedge \{H \vee W\} \\
&= \{O \wedge T \wedge H \wedge W\}
\end{aligned}
$$

Therefore, by Theorem 21, only one 0.3-decision reduct is obtained: $B = \{O, T, H, W\}$.

Now, we will perform the same calculations but with $\delta = 0.2$. In this case, the 0.2-discernibility matrix $O_{0.2}$ is:

$$
\begin{pmatrix}
\varnothing & & & & & \\
\{W\} & \varnothing & & & & \\
\{O\} & \{O, W\} & \varnothing & & & \\
\{O, T\} & \{O, T, W\} & \{T\} & \varnothing & & \\
\{O, T\} & \{O, T, W\} & \{T\} & \varnothing & \varnothing & \\
\{O, T, W\} & \{O, T\} & \{T, W\} & \{W\} & \{W\} & \varnothing
\end{pmatrix}
$$

Hence, the unidimensional 0.2-discernibility function is

$$
\begin{aligned}
\tau^{uni} &= \{W\} \wedge \{O\} \wedge \{O \vee T\} \wedge \{O \vee W\} \\
&\wedge \{O \vee T \vee W\} \wedge \{T \vee W\} \wedge \{T\} \\
&= \{O \wedge T \wedge W\}
\end{aligned}
$$

Hence, only one 0.2-decision reduct is obtained: $B = \{O, T, W\}$.

4. δ-information bireducts

This section is focused on introducing and characterizing the bireducts considering the new framework with similarities.

The following definition is the natural extension of the discernibility function expression to δ-information bireducts.

Definition 23 *Let $\mathbb{A} = (U, \mathcal{A})$ be a information system, the conjunctive normal form*

$$
\bigwedge \{i \vee j \bigvee \{a \in \mathcal{A} \mid E_a(a(i), a(j)) < \delta\} \mid i, j \in U\}
$$

where the elements of U and $\mathcal{A}$ are the propositional symbols of the language, is called the bidimensional δ-discernibility function *and it is denoted as $\tau_{\mathcal{A}}^{bi}$.*

The following theorem characterizes the δ-information bireducts.

Theorem 24 *Given an information system $\mathbb{A} = (U, \mathcal{A})$, an arbitrary pair (B, X), $B \subseteq \mathcal{A}$, $X \subseteq U$, is a δ-information bireduct if and only if the cube $\bigwedge_{b \in B} b \wedge \bigwedge_{i \notin X} i$ is a cube in the RDNF of $\tau_{\mathbb{A}}^{bir}$.*

Theorem 24 will be applied in the following example to the weather information system given in the previous section.

Example 25 *From the information system given in Example 9, we will consider the thresholds $\delta = 0.3$ and $\delta = 0.2$ in order to obtain the corresponding δ-information bireducts.*

First of all, we compute the bidimensional 0.3-discernibility function, by Definition 23, adding the corresponding attributes to the 0.3-discernibility matrix $O_{0.3}$:

$$
\begin{aligned}
\tau^{bi} &= \{1 \vee 2 \vee W\} \wedge \{1 \vee 3 \vee O\} \wedge \{1 \vee 4 \vee O \vee T\} \\
&\wedge \{1 \vee 5 \vee O \vee T \vee H\} \\
&\wedge \{1 \vee 6 \vee O \vee T \vee H \vee W\} \\
&\wedge \{2 \vee 3 \vee O \vee W\} \wedge \{2 \vee 4 \vee O \vee T \vee W\} \\
&\wedge \{2 \vee 5 \vee O \vee T \vee H \vee W\} \\
&\wedge \{2 \vee 6 \vee O \vee T \vee H\} \\
&\wedge \{3 \vee 4 \vee T\} \wedge \{3 \vee 5 \vee T \vee H\} \\
&\wedge \{3 \vee 6 \vee T \vee H \vee W\} \wedge \{4 \vee 5 \vee H\} \\
&\wedge \{4 \vee 6 \vee H \vee W\} \wedge \{5 \vee 6 \vee W\}
\end{aligned}
$$

From this CNF we compute the RDNF. After that, 80 cubes are obtained, which provides 80 bidimensional 0.3-information bireducts. For example, some of them are:

$$
\begin{aligned}
(B_1, X_1) &= (\{H, W\}, \{1, 3\}) \\
(B_2, X_2) &= (\{O, W\}, \{3, 5\}) \\
(B_3, X_3) &= (\{O, T, H, W\}, \{\}) \\
(B_4, X_4) &= (\{O, T, W\}, \{4\}) \\
(B_5, X_5) &= (\{W\}, \{1, 2, 3, 4\}) \\
(B_6, X_6) &= (\{T, W\}, \{2, 3, 5\}) \\
(B_7, X_7) &= (\{O, T, H\}, \{2, 5\}) \\
(B_8, X_8) &= (\{\}, \{1, 2, 3, 4, 6\})
\end{aligned}
$$

Taken into account the threshold $\delta = 0.2$ and the 0.2-discernibility matrix $=_{0.2}$, the bidimensional 0.2-discernibility function is

$$
\begin{aligned}
\tau^{bi} = \;& \{1 \vee 2 \vee W\} \wedge \{1 \vee 3 \vee W\} \wedge \{1 \vee 4 \vee O \vee T\} \\
& \wedge \{1 \vee 5 \vee O \vee T\} \wedge \{1 \vee 6 \vee O \vee T \vee W\} \\
& \wedge \{2 \vee 3 \vee O \vee W\} \wedge \{2 \vee 4 \vee O \vee T \vee W\} \\
& \wedge \{2 \vee 5 \vee O \vee T \vee W\} \wedge \{2 \vee 6 \vee O \vee T\} \\
& \wedge \{3 \vee 4 \vee T\} \wedge \{3 \vee 5 \vee T\} \wedge \{3 \vee 6 \vee T \vee W\} \\
& \wedge \{4 \vee 6 \vee W\} \wedge \{5 \vee 6 \vee W\}
\end{aligned}
$$

The RDNF has 23 cubes, which provides the corresponding 23 bidimensional 0.2-decision bireducts, such as:

$$
\begin{aligned}
(B_1, X_1) &= (\{T, W\}, \{\}) \\
(B_2, X_2) &= (\{O, W\}, \{3\}) \\
(B_3, X_3) &= (\{\}, \{1, 2, 3, 6\}) \\
(B_4, X_4) &= (\{W\}, \{1, 2, 3\}) \\
(B_5, X_5) &= (\{O\}, \{2, 3, 6\}) \\
(B_6, X_6) &= (\{O, T\}, \{1, 6\}) \\
(B_7, X_7) &= (\{O, W\}, \{4, 5\}) \\
(B_8, X_8) &= (\{\}, \{1, 3, 4, 6\})
\end{aligned}
$$

5. Conclusions and future work

We have studied the reducts and bireducts in the classic environment of rough set theory and considering similarity relations. We have generalized the discernibility function, from which we could get the reducts and bireducts in these environments. The inclusion of the similarity relations in theory provides a greater flexibility in these environments, dramatically increasing the range of possible applications.

As future work, we will extend the theory to obtain bireducts in FCA and fuzzy environments, such as in fuzzy rough sets. Moreover, we will study in depth in the relation between concept lattice reduction and rough set reduction considering similarity relations and in the general fuzzy case, considering the ideas given in [6, 10, 12, 24]. Furthermore, we apply the theory developed in both theories to practical cases.

References

[1] J. Bazan, A. Skowron, D. Ślęzak, and J. Wróblewski. Searching for the complex decision reducts: The case study of the survival analysis. In N. Zhong, Z. Ra, S. Tsumoto, and E. Suzuki, editors, *Foundations of Intelligent Systems*, volume 2871 of *Lecture Notes in Computer Science*, pages 160–168. Springer Berlin Heidelberg, 2003.

[2] Y. Chen and Y. Yao. A multiview approach for intelligent data analysis based on data operators. *Information Sciences*, 178(1):1–20, 2008.

[3] C. Cornelis, R. Jensen, G. Hurtado, and D. Ślęzak. Attribute selection with fuzzy decision reducts. *Information Sciences*, 180:209–224, 2010.

[4] I. Düntsch and G. Gediga. Approximation operators in qualitative data analysis. In *Theory and Applications of Relational Structures as Knowledge Instruments*, pages 214–230, 2003.

[5] L. Fariñas del Cerro and H. Prade. Rough sets, twofold fuzzy sets and modal logic—fuzziness in indiscernibility and partial information. In A. D. Nola and A. Ventre, editors, *The Mathematics of Fuzzy Systems*, pages 103–120. Verlag TUV Rheinland, 1986.

[6] B. Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundation.* Springer Verlag, 1999.

[7] A. Janusz, D. Ślęzak, and H. S. Nguyen. Unsupervised similarity learning from textual data. *Fundamenta Informaticae*, Volume 119:319 – 336, 01 2012.

[8] A. Krętowski and J. Stepaniuk. Selection of objects and attributes a tolerance rough set approach. In *Proceedings of the Poster Session of Ninth International Symposium on Methodologies for Intelligent Systems*, pages 169–180, Zakopane, Poland, July 10-13 1996.

[9] J. Li, C. Mei, and Y. Lv. Incomplete decision contexts: Approximate concept construction, rule acquisition and knowledge reduction. *International Journal of Approximate Reasoning*, 54(1):149–165, 2013.

[10] M. Liu, M. Shao, W. Zhang, and C. Wu. Reduction method for concept lattices based on rough set theory and its application. *Computers & Mathematics with Applications*, 53(9):1390–1410, 2007.

[11] N. Mac Parthalain and R. Jensen. Simultaneous feature and instance selection using fuzzy-rough bireducts. In *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*, pages 1–8, July 2013.

[12] J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.

[13] N. N. Morsi and M. M. Yakout. Axiomatics

for fuzzy rough sets. *Fuzzy sets and Systems*, 100:327–342, 1998.

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

[15] S. Nanda and S. Majumdar. Fuzzy rough sets. *Fuzzy sets and Systems*, 45:157–160, 1992.

[16] Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.

[17] A. M. Radzikowska and E. E. Kerre. A comparative study of fuzzy rough sets. *Fuzzy Sets and Systems*, 126(2):137–155, 2002.

[18] A. Skowron and C. Rauszer. The discernibility matrices and functions in information systems. In R. Słowiński, editor, *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory*, pages 331–362. Kluwer Academic Publishers, 1992.

[19] A. Skowron and J. Stepaniuk. Tolerance approximation spaces. *Fundamenta Informaticae*, 27(2-3):245–253, 1996.

[20] D. Ślęzak and A. Janusz. Ensembles of bireducts: Towards robust classification and simple representation. In T.-h. Kim, H. Adeli, D. Ślęzak, F. Sandnes, X. Song, K.-i. Chung, and K. Arnett, editors, *Future Generation Information Technology*, volume 7105 of *Lecture Notes in Computer Science*, pages 64–77. Springer Berlin Heidelberg, 2011.

[21] S. Stawicki and D. Ślęzak. Recent advances in decision bireducts: Complexity, heuristics and streams. *Lecture Notes in Computer Science*, 8171:200–212, 2013.

[22] J. Stepaniuk and A. Krętowski. Decision system based on tolerance rough sets. In *Proceedings of the Fourth International Workshop on Intelligent Information Systems*, pages 62–73, Augustow, Poland, June 5-9 1995. Institute of Computer Science, Polish Academy of Sciences.

[23] X. Wang and W. Zhang. Relations of attribute reduction between object and property oriented concept lattices. *Knowledge-Based Systems*, 21:398–403, 2008.

[24] L. Wei and J.-J. Qi. Relation between concept lattice reduction and rough set reduction. *Knowledge-Based Systems*, 23(8):934–938, 2010.

[25] Y. Yao and Y. Zhao. Attribute reduction in decision-theoretic rough set models. *Information Sciences*, 178(17):3356–3373, 2008.