Efficiency of fuzzy decision algorithms based on the certainty of decision rules*

Fernando Chacón-Gómez[0000−0001−9710−7813], M. Eugenia Cornejo[0000−0002−8230−0044], Jesús Medina[0000−0002−3931−5873]

Department of Mathematics, University of Cádiz, Cádiz, Spain
{fernando.chacon,mariaeugenia.cornejo,jesus.medina}@uca.es

Abstract. Decision algorithms ensure an appropriate extraction of information from datasets by means of decision rules. These algorithms are analyzed by the notion of efficiency, which focuses on the objects correctly classified by using only the most representative decision rules. This paper presents a novel notion of efficiency in the fuzzy framework based on the certainty of the decision rules, in order to quantify the quality of classification of the studied algorithm.

Keywords: Fuzzy Rough Set Theory, Decision Rules, Decision Algorithm, Efficiency.

1 Introduction

Pawlak introduced Rough Set Theory (RST) [16,17] in the eighties as a mathematical tool to analyze relational datasets, which are interpreted as decision tables, containing imprecise or incomplete information. Decision rules [10,19] are considered in this framework to simplify the extraction of information and its interpretation by means of several relevance indicators, such as the strength and the certainty. It is important that the set of rules used to analyze a dataset allows us to extract information in a suitably way.

With this purpose, the notion of decision algorithm arises as a set of decision rules that satisfy some requirements, such as the provision of non-redundancy information and the preservation of the consistency of the decision table. Pawlak also introduced the notion of efficiency [18] to analyze decision algorithms. This notion provides the proportion of objects that satisfy the most representative rule for each different antecedent, so that higher values of efficiency correspond to better decision algorithms. In addition, this notion is especially useful for comparing different decision algorithms and extracting conclusions.

On the other hand, Fuzzy Rough Set Theory (FRST) [8,13,14,15] is a natural extension of RST to the fuzzy framework based on the philosophy of fuzzy

* Partially supported by the 2014–2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in projects PID2019-108991GB-I00 and PID2022-137620NB-I00, and with the Ecological and Digital Transition Projects 2021 of the Ministry of Science and Innovation in project TED2021-129748B-I00.

sets [22]. This mathematical tool has been extensively studied, directly, such as in [5,6,9,21] and from other perspectives [1,11,12,20], due to its advantages in managing uncertain and inconsistent information by offering a great flexibility to the corresponding analysis. However, many developments are still required. For example, the notions of algorithm of decision rules and efficiency of an algorithm [18] were not completely accommodated to the fuzzy setting. Recent advances in these extensions have been made in [2,3,4], but it remains different improvements concerning the notion of efficiency.

The classical definition of efficiency was based on the strength of the decision rules, and a first extension was introduced in [2]. This paper focuses on the introduction of a new definition of efficiency in FRST to analyze decision algorithms from a different point of view, that is, based on the certainty of the decision rules. Moreover, we will present some interesting properties satisfied by this notion, which will simplify its computation under certain hypothesis and analyze the variability of the possible results. The new definition and the properties are illustrated in a final example.

2 Basic notions in FRST

This section is devoted to recalling important notions of FRST [4] for computing the efficiency of decision algorithms. To begin with, decision tables are introduced which represent datasets in this framework.

Definition 1. *Let* U *and* $\mathcal{A}$ *be non-empty sets of objects and attributes, respectively. A* decision table *is a tuple* $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ *such that* $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ *with* $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, *where* V_a *is the set of values associated with the attribute* a *over* U, *and* $\overline{\mathcal{A}_d} = \{\bar{a} \mid a \in \mathcal{A}_d, \bar{a} \colon U \to V_a\}$. *In this case, the attributes of* $\mathcal{A}$ *are called* condition attributes *and* d *is called* decision attribute.

Next, we present the notions of formula, degree of satisfiability to a given formula, conjunction/disjunction of formulas and decision rule, in order to analyze the information contained in decision tables more easily. Notice that, the degree of satisfiability to a given formula is based on separable tolerance relations [23].

Definition 2. *Let* $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ *be a decision table,* $B \subseteq \mathcal{A}_d$, $C \subseteq \mathcal{A}$ *and* $T = \{T_a \colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}_d\}$ *be a family of separable* $[0,1]$*-fuzzy tolerance relations, that is, each mapping* T_a *is a symmetrical fuzzy relationship satisfying that* $T_a(v,w) = 1$ *if and only if* $v = w$.

- *The* set of formulas associated with B, *denoted as* $For(B)$, *is built from attribute-value pairs* (a, v), *where* $a \in B$ *and* $v \in V_a$, *by means of the conjunction and disjunction logical connectives,* $\wedge$ *and* $\vee$, *respectively.*
- *The mapping* $\|\cdot\|_S^T \colon For(B) \to [0,1]^U$ *inductively defined as*

$$\|\Phi\|_S^T(x) = T_a(\bar{a}(x), v)$$

for each $x \in U$ *and* $\Phi = (a, v)$, *with* $a \in B$ *and* $v \in V_a$, *is the* degree of satisfiability to the formula Φ of the object x, *through the relationships* *between the values of the attributes in the object x and the values of the attributes in the formula Φ.*

- *For every $\Phi, \Psi \in For(B)$, the* conjunction and disjunction of formulas *is defined, for each $x \in U$, as follows:*

$$\|\Phi \wedge \Psi\|_S^T(x) = \min\{\|\Phi\|_S^T(x), \|\Psi\|_S^T(x)\}$$
$$\|\Phi \vee \Psi\|_S^T(x) = \max\{\|\Phi\|_S^T(x), \|\Psi\|_S^T(x)\}$$

- *A* decision rule *in S is an expression $\Phi \to \Psi$, where $\Phi \in For(C)$ is the antecedent of the decision rule and $\Psi \in For(\{d\})$ is the consequent of the decision rule.*

Decision rules allow to summarize decision tables in logical terms and they are described by several fuzzy relevance indicators from different perspectives [4]. In particular, we recall the T-strength and the T-certainty, which are defined from the classical cardinal and the fuzzy cardinal [7].

Definition 3. *Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\Phi \to \Psi$ be a decision rule in S and $T = \{T_a \colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}_d\}$ be a family of separable $[0,1]$-fuzzy tolerance relations. We call:*

- T-strength *of the decision rule $\Phi \to \Psi$ to the value:*

$$\sigma_S^T(\Phi, \Psi) = \frac{card_F(\|\Phi \wedge \Psi\|_S^T)}{card(U)}$$

- T-certainty *of the decision rule $\Phi \to \Psi$ to the value:*

$$cer_S^T(\Phi, \Psi) = \frac{card_F(\|\Phi \wedge \Psi\|_S^T)}{card_F(\|\Phi\|_S^T)}$$

where $card_F(\cdot)$ denotes the cardinal of a fuzzy set and $card(\cdot)$ denotes the cardinal of a classical set.

The T-strength of a decision rule provides its representativeness in the decision table under consideration, while the T-certainty represents how much the antecedent implies the consequent. The following notion is useful to compare pairs of formulas and to analyze their similarity according to a threshold.

Definition 4. *Given a decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$, a family of separable $[0,1]$-fuzzy tolerance relations $T = \{T_a \colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}_d\}$ and $\Phi, \Phi' \in For(\mathcal{A}_d)$, such that*

$$\Phi = (a_1, v_1) \wedge \ldots \wedge (a_n, v_n) \quad \text{and} \quad \Phi' = (a'_1, w_1) \wedge \ldots \wedge (a'_m, w_m)$$

we define the F-indiscernibility relation *as a separable $[0,1]$-fuzzy tolerance relation $R_{Fd} \colon For(\mathcal{A}_d) \times For(\mathcal{A}_d) \to [0,1]$ given by*

$$R_{Fd}(\Phi, \Phi') = \begin{cases} \bigwedge\limits_{i \in \{1,\ldots,n\}} T_{a_i}(v_i, w_i) & \text{if } n = m \text{ and } a_i = a'_i \text{ for each } i \in \{1, \ldots, n\} \\ 0 & \text{otherwise} \end{cases}$$

Given $\alpha \in [0,1]$*, we define the* R_{Fd}-α-block of $\Phi \in For(\mathcal{A}_d)$ *as follows:*

$$[\Phi]_\alpha = \{\Phi' \in For(\mathcal{A}_d) \mid \alpha \leq R_{Fd}(\Phi, \Phi')\}$$

If $\Phi' \in [\Phi]_\alpha$*, then we will say that* Φ and Φ' are R_{Fd}-α-related*.*

R_{Fd}-α-blocks summarize the relationship between formulas. Moreover, they depend on a threshold α, which is chosen by the user according to his/her necessities. These blocks are also useful to obtain those decision rules without contradictions between them, which are the most reliable ones in a given set of decision rules.

Definition 5. *Let* $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ *be a decision table,* $\alpha \in (0,1]$*,* $R_{Fd}\colon For(\mathcal{A}_d) \times For(\mathcal{A}_d) \to [0,1]$ *be a* F*-indiscernibility relation and* $Dec(S) = \{\Phi_i \to \Psi_i \mid i \in \{1,\dots,m\}\}$ *a set of decision rules in* S*. The set of* α-consistent decision rules *is defined as follows:*

$$Dec^+_\alpha(S) = \{\Phi \to \Psi \in Dec(S) \mid \textit{if for each } \Phi' \to \Psi' \in Dec(S) \textit{ such that } \Phi' \in [\Phi]_\alpha \textit{ then } \Psi' \in [\Psi]_\alpha\}$$

Before presenting decision algorithms in FRST, it is necessary to introduce the fuzzy positive region. This notion is defined from indiscernibility relations and the multi-adjoint property-oriented frame, which were given in [6].

Definition 6. *Let* $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ *be a decision table and* R_d *be a d-indiscernibility boolean relation. The* multi-adjoint fuzzy $\mathcal{A}$-positive region *is defined, for each* $y \in U$*, as:*

$$POS^f_{\mathcal{A}}(y) = \inf\{R_d(y,x) \nwarrow_{\tau(x,y)} R_{\mathcal{A}}(x,y) \mid x \in U\}$$

where $\nwarrow_{\tau(x,y)}$ *is the left residuated fuzzy implication of* $\&_{\tau(x,y)}$ *associated with the pair of objects* x, y*.*

Now, we recall the notion of decision algorithm [4], which is a set of decision rules satisfying some requirements to ensure that they collect the most significant and non-redundant information in the table.

Definition 7. *Let* $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ *be a decision table,* $\alpha_1, \alpha_2, \alpha_4 \in [0,1)$*,* $\alpha_3 \in (0, card(U)]$*,* $\alpha \in (0,1]$*,* $T = \{T_a\colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}_d\}$ *be a family of separable* $[0,1]$*-fuzzy tolerance relations and* $Dec(S) = \{\Phi_i \to \Psi_i \mid i \in \{1,\dots,m\}\}$*, with* $m \geq 2$*, be a set of decision rules in* S*. We say that:*

1. $Dec(S)$ *is a set of* $\alpha_1\alpha_2$-pairwise mutually exclusive (independent) *decision rules, if each pair of decision rules* $\Phi \to \Psi, \Phi' \to \Psi' \in Dec(S)$ *satisfies that* $\Phi = \Phi'$ *or* $\|\Phi \wedge \Phi'\|^T_S(x) \leq \alpha_1$ *and* $\Psi = \Psi'$ *or* $\|\Psi \wedge \Psi'\|^T_S(x) \leq \alpha_2$*, for all* $x \in U$*.*
2. $Dec(S)$ covers U*, if* $card_F(\|\bigvee_{i=1}^{m} \Phi_i\|^T_S) = card_F(\|\bigvee_{i=1}^{m} \Psi_i\|^T_S) = card(U)$*.*
3. *The decision rule* $\Phi \to \Psi \in Dec(S)$ *is* α_3-admissible in S *if* $\alpha_3 \leq supp^T_S(\Phi, \Psi)$*.*

4. *Dec*(S) preserves the α-consistency of S with a degree α_4 *if the next inequality holds for each* $x \in U$*:*

$$\left| POS^f_{\mathcal{A}}(x) - \bigvee_{\Phi \to \Psi \in Dec^+_\alpha(S)} \|\Phi\|^T_S(x) \right| \leq \alpha_4$$

The set of decision rules $Dec(S)$ *satisfying the previous properties for the values* $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ *is called* $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$-decision algorithm in S *and it is denoted as* $DA_T(S)$*.*

More information about decision algorithms in FRST can be consulted in [4]. Finally, we recall the efficiency of decision algorithms in FRST, which generalizes the classical notion of efficiency introduced by Pawlak [18]. This definition was introduced, together with different properties, in [2,3].

Definition 8. *Let* $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ *be a decision table and* $DA_T(S)$ *be a* $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$*-decision algorithm. Given* $\varepsilon \in [0, 1]$*, we call* ε-efficiency *of* $DA_T(S)$ *to the number*

$$\eta^\varepsilon(DA_T(S)) = \sum_{\{\Phi \in For(\mathcal{A}) \mid \Phi \to \Psi \in DA_T(S)\}} \eta^\varepsilon_\Phi(DA_T(S))$$

where

$$\eta^\varepsilon_\Phi(DA_T(S)) = \max\{\sigma^T_S(\Phi', \Psi') \mid \Phi' \to \Psi' \in DA_T(S), \Phi' \in [\Phi]_\varepsilon\}$$

The following example analyzes a decision table by using the basic notions in FRST introduced in this section.

Example 1. Consider the decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ represented in Table 1, whose set of objects is $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, set of condition attributes is $\mathcal{A} = \{a_1, a_2, a_3\}$ and $V_a = [0, 1]$ for all $a \in \mathcal{A}_d$.

Table 1: Decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 1.

	a_1	a_2	a_3	d
x_1	0.34	0.31	0.75	0
x_2	0.21	0.71	0.5	1
x_3	0.52	0.92	1	0
x_4	0.85	0.65	1	1
x_5	0.43	0.89	0.5	0
x_6	0.21	0.47	0.25	1
x_7	0.09	0.93	0.25	0

In order to preserve all the information contained in S, we extract a decision rule from each object in Table 1, by considering the conjunction of all the condition attributes. As a consequence, we obtain the set of decision rules

$Dec(S) = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7\}$ given below:

$$r_1 : (a_1, 0.34) \wedge (a_2, 0.31) \wedge (a_3, 0.75) \rightarrow (d, 0)$$
$$r_2 : (a_1, 0.21) \wedge (a_2, 0.71) \wedge (a_3, 0.5) \rightarrow (d, 1)$$
$$r_3 : (a_1, 0.52) \wedge (a_2, 0.92) \wedge (a_3, 1) \rightarrow (d, 0)$$
$$r_4 : (a_1, 0.85) \wedge (a_2, 0.65) \wedge (a_3, 1) \rightarrow (d, 1)$$
$$r_5 : (a_1, 0.43) \wedge (a_2, 0.89) \wedge (a_3, 0.5) \rightarrow (d, 0)$$
$$r_6 : (a_1, 0.21) \wedge (a_2, 0.47) \wedge (a_3, 0.25) \rightarrow (d, 1)$$
$$r_7 : (a_1, 0.09) \wedge (a_2, 0.93) \wedge (a_3, 0.25) \rightarrow (d, 0)$$

which are denoted as $r_i : \Phi_i \rightarrow \Psi_i$, for all $i \in \{1, \ldots, 7\}$.

As was shown in [4], taking $\alpha = 0.75$, $\alpha_1 \geq 0.78$, $\alpha_2 \geq 0$, $\alpha_3 \leq 1.61$ and $\alpha_4 \geq 0.39$, we deduce that $Dec(S)$ is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$-decision algorithm. From now on, consider $DA_T(S) = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7\}$.

Finally, in order to compute the ε-efficiency of $DA_T(S)$ attending to Definition 8, we set three different thresholds ε, corresponding to low, medium and high values, respectively. In particular, we choose the values 0.25, 0.8 and 0.67. Following the procedure exposed in [2], the different ε-efficiencies of $DA_T(S)$ are computed, whose results are presented in Table 2. □

Table 2: Different ε-efficiencies of $DA_T(S)$.

Threshold	$\eta^\varepsilon(DA_T(S))$
$\varepsilon = 0.25$	2.59
$\varepsilon = 0.8$	2.14
$\varepsilon = 0.67$	2.3

One of the disadvantages of Definition 8 is that the ε-efficiency is not bounded by 1, so it may be difficult to interpret the obtained values. Moreover, it is intuitive to think that the certainty of the selected decision rules is also important to measure the efficiency of an algorithm. These issues are considered in the next section by means of the introduction of a novel notion of efficiency taking into account the certainty of decision rules, together with the inclusion of its main properties showing the relevance of this notion as indicator of a decision algorithm.

3 Efficiency of decision algorithms based on the certainty

Now, we will study the efficiency of decision algorithms in FRST from a different perspective. For that, we introduce an alternative notion to Definition 8, which considers the T-certainty of the rules instead of the T-strength.

Definition 9. *Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $DA_T(S)$ be a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$-decision algorithm. Given $\varepsilon \in [0, 1]$, we call c_ε-*efficiency *of*

$DA_T(S)$ to the number

$$\eta^{c_\varepsilon}(DA_T(S)) = \sum_{\{\Phi \in For(\mathcal{A}) \mid \Phi \to \Psi \in DA_T(S)\}} \frac{\eta_\Phi^{c_\varepsilon}(DA_T(S))}{card(\{\Phi \in For(\mathcal{A}) \mid \Phi \to \Psi \in DA_T(S)\})}$$

where

$$\eta_\Phi^{c_\varepsilon}(DA_T(S)) = \max\{cer_S^T(\Phi', \Psi') \mid \Phi' \to \Psi' \in DA_T(S), \Phi' \in [\Phi]_\varepsilon\}$$

Although Definition 9 does not generalize the classical notion of efficiency introduced by Pawlak [18], it also measures the efficiency of decision algorithms, since it provides their quality of classification by means of the average of the T-certainties under consideration. Next result presents some interesting properties deduced from Definition 9.

Proposition 1. *Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $DA_T(S) = \{\Phi_i \to \Psi_i \mid i \in \{1, \ldots, m\}\}$, with $m \geq 2$, be a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$-decision algorithm. The following properties hold:*

1. *The c_ε-efficiency of $DA_T(S)$ is decreasing in ε, that is, given $\varepsilon_1, \varepsilon_2 \in [0, 1]$, if $\varepsilon_1 \leq \varepsilon_2$ then*
$$\eta^{c_{\varepsilon_2}}(DA_T(S)) \leq \eta^{c_{\varepsilon_1}}(DA_T(S))$$
2. *Let $\Phi_j \to \Psi_j \in DA_T(S)$ such that*
$$cer_S^T(\Phi_j, \Psi_j) = \max\{cer_S^T(\Phi, \Psi) \mid \Phi \to \Psi \in DA_T(S)\}$$
If $\varepsilon \leq \min\{R_{Fd}(\Phi_i, \Phi_j) \mid i \in \{1, \ldots, m\}\}$, then the c_ε-efficiency of $DA_T(S)$ is given as
$$\eta^{c_\varepsilon}(DA_T(S)) = cer_S^T(\Phi_j, \Psi_j)$$
3. *If $\max\{R_{Fd}(\Phi_i, \Phi_j) \mid i, j \in \{1, \ldots, m\}, i \neq j\} < \varepsilon$, then the c_ε-efficiency of $DA_T(S)$ is given as*
$$\eta^{c_\varepsilon}(DA_T(S)) = \frac{\sum_{\Phi \to \Psi \in DA_T(S)} cer_S^T(\Phi, \Psi)}{card(\{\Phi \in For(\mathcal{A}) \mid \Phi \to \Psi \in DA_T(S)\})}$$

Proposition 1(1) ensures that the c_ε-efficiency of a decision algorithm $DA_T(S)$ is decreasing in the threshold ε. This is consistent with the fact that, if we decrease the threshold more antecedents are considered in the corresponding classes and a greater certainty is considered per antecedent. Hence, a good balance between the threshold ε and the c_ε-efficiency is required. On the other hand, Proposition 1(2) exposes the case that, due to the choice of the threshold ε, only the decision rule with the greatest T-certainty is considered, since its antecedent is R_{Fd}-ε-related to the antecedents of the rest of rules. Finally, Proposition 1(3) focuses on the case in which none of the antecedents of the rules in $DA_T(S)$ are R_{Fd}-ε-related. As a result, all the rules are needed to compute the c_ε-efficiency of $DA_T(S)$. Moreover, notice that, from the three items of Proposition 1, it is possible to determine the range of all values that the c_ε-efficiency of any decision algorithm can take.

Corollary 1. *Let* $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ *be a decision table and* $DA_T(S) = \{\Phi_i \to \Psi_i \mid i \in \{1, \ldots, m\}\}$*, with* $m \geq 2$*, be a* $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$*-decision algorithm. If* $\Phi_i \neq \Phi_j$ *for each* $i, j \in \{1, \ldots, m\}$ *with* $i \neq j$*, then for all* $\varepsilon \in [0,1]$*:*

$$\frac{\displaystyle\sum_{\Phi \to \Psi \in DA_T(S)} cer_S^T(\Phi, \Psi)}{card(\{\Phi \in For(\mathcal{A}) \mid \Phi \to \Psi \in DA_T(S)\})} \leq \eta^{c_\varepsilon}(DA_T(S)) \leq \max\{cer_S^T(\Phi, \Psi) \mid \Phi \to \Psi \in DA_T(S)\}$$

The previous result provides an interval useful to detect the influence of the threshold ε in the c_ε-efficiency of a decision algorithm $DA_T(S)$. For instance, if the certainties of all the decision rules in $DA_T(S)$ are similar, then that interval is small. Consequently, in this case, the c_ε-efficiency does not depend much on the fixed threshold ε. Moreover, if all the rules are certain, then the efficiency is 1, independently of the threshold. In particular, the c_ε-efficiency of any decision algorithm is a value of the unit interval for all $\varepsilon \in [0,1]$, which allows us to interpret the algorithms more easily than applying Definition 8.

Now, we continue analyzing the decision algorithm given in Example 1 by taking advantage of Definition 9, Proposition 1 and Corollary 1.

Example 2. First of all, we will calculate the T-certainty of each decision rule in the decision algorithm $DA_T(S)$ given in Example 1, in order to determine the c_ε-efficiency of $DA_T(S)$. With this purpose, we consider the family $T = \{T_a \colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}_d\}$ defined as $T_a(v, w) = 1 - |v - w|$, for each $a \in \mathcal{A}_d$ and $v, w \in V_a$, in Definition 3. The obtained results are presented in Table 3.

Table 3: T-certainty of decision rules in $DA_T(S)$.

Rule	cer_S^T
r_1	0.58
r_2	0.45
r_3	0.6
r_4	0.46
r_5	0.58
r_6	0.52
r_7	0.59

Furthermore, it is also necessary to compute the relation between each pair of antecedents in $DA_T(S)$, applying Definition 4. By using the aforementioned family T, we obtain the results exposed in Table 4.

Now, we will consider the same values of the threshold ε as in Example 8 in order to compute different c_ε-efficiencies of $DA_T(S)$.

- To begin with, we compute the $c_{0.25}$-efficiency of $DA_T(S)$. Notice that, from Table 3, the decision rule with the greatest T-certainty is r_3. Moreover, we have that $[\Phi_3]_{0.25} = \{\Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_5, \Phi_6, \Phi_7\}$ from Table 4 and Definition 4. Hence, by Proposition 1(2) we deduce that:

$$\eta^{c_{0.25}}(DA_T(S)) = cer_S^T(\Phi_3, \Psi_3) = 0.6$$

Table 4: Relation between each pair of antecedents of decision rules in $DA_T(S)$.

R_{Fd}	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7
Φ_1	1	0.6	0.39	0.49	0.42	0.5	0.38
Φ_2	0.6	1	0.5	0.36	0.78	0.75	0.75
Φ_3	0.39	0.5	1	0.67	0.5	0.25	0.25
Φ_4	0.49	0.36	0.67	1	0.5	0.25	0.24
Φ_5	0.42	0.78	0.5	0.5	1	0.58	0.66
Φ_6	0.5	0.75	0.25	0.25	0.58	1	0.54
Φ_7	0.38	0.75	0.25	0.24	0.66	0.54	1

- We continue the study by using the threshold $\varepsilon = 0.8$. Since $[\Phi_i]_{0.8} = \{\Phi_i\}$ attending to Table 4, for all $i \in \{1, \ldots, 7\}$, by Proposition 1(3) we obtain that:

$$\eta^{c_{0.8}}(DA_T(S)) = \frac{\displaystyle\sum_{\Phi\to\Psi\in DA_T(S)} cer_S^T(\Phi,\Psi)}{\mathrm{card}(\{\Phi \in For(\mathcal{A}) \mid \Phi \to \Psi \in DA_T(S)\})}$$
$$= \frac{0.58 + 0.45 + 0.6 + 0.46 + 0.58 + 0.52 + 0.59}{7} = 0.54$$

- In order to compute the $c_{0.67}$-efficiency, it is necessary to apply Definition 9. Taking into account Definition 4 and Table 4, we obtain the following R_{Fd}-0.67-blocks:

$$[\Phi_1]_{0.67} = \{\Phi_1\} \quad [\Phi_4]_{0.67} = \{\Phi_3, \Phi_4\} \quad [\Phi_6]_{0.67} = \{\Phi_2, \Phi_6\}$$
$$[\Phi_2]_{0.67} = \{\Phi_2, \Phi_5, \Phi_6, \Phi_7\} \quad [\Phi_5]_{0.67} = \{\Phi_2, \Phi_5\} \quad [\Phi_7]_{0.67} = \{\Phi_2, \Phi_7\}$$
$$[\Phi_3]_{0.67} = \{\Phi_3, \Phi_4\}$$

Thus, by Definition 9 we obtain that:

$$\eta_{\Phi_1}^{c_{0.67}}(DA_T(S)) = cer_S^T(\Phi_1, \Psi_1) = 0.58$$
$$\eta_{\Phi_2}^{c_{0.67}}(DA_T(S)) = \max\{cer_S^T(\Phi_2, \Psi_2), cer_S^T(\Phi_5, \Psi_5), cer_S^T(\Phi_6, \Psi_6), cer_S^T(\Phi_7, \Psi_7)\}$$
$$= cer_S^T(\Phi_7, \Psi_7) = 0.59$$
$$\eta_{\Phi_3}^{c_{0.67}}(DA_T(S)) = \max\{cer_S^T(\Phi_3, \Psi_3), cer_S^T(\Phi_4, \Psi_4)\} = cer_S^T(\Phi_3, \Psi_3) = 0.6$$
$$\eta_{\Phi_4}^{c_{0.67}}(DA_T(S)) = \max\{cer_S^T(\Phi_3, \Psi_3), cer_S^T(\Phi_4, \Psi_4)\} = cer_S^T(\Phi_3, \Psi_3) = 0.6$$
$$\eta_{\Phi_5}^{c_{0.67}}(DA_T(S)) = \max\{cer_S^T(\Phi_2, \Psi_2), cer_S^T(\Phi_5, \Psi_5)\} = cer_S^T(\Phi_5, \Psi_5) = 0.58$$
$$\eta_{\Phi_6}^{c_{0.67}}(DA_T(S)) = \max\{cer_S^T(\Phi_2, \Psi_2), cer_S^T(\Phi_6, \Psi_6)\} = cer_S^T(\Phi_6, \Psi_6) = 0.52$$
$$\eta_{\Phi_7}^{c_{0.67}}(DA_T(S)) = \max\{cer_S^T(\Phi_2, \Psi_2), cer_S^T(\Phi_7, \Psi_7)\} = cer_S^T(\Phi_7, \Psi_7) = 0.59$$

As a result, the $c_{0.67}$-efficiency of $DA_T(S)$ is:

$$\eta^{c_{0.67}}(DA_T(S)) = \frac{\displaystyle\sum_{\{\Phi\in For(\mathcal{A})\mid \Phi\to\Psi\in DA_T(S)\}} \eta_{\Phi}^{c_{0.67}}(DA_T(S))}{\text{card}(\{\Phi \in For(\mathcal{A}) \mid \Phi \to \Psi \in DA_T(S)\})}$$

$$= \frac{0.58 + 0.59 + 0.6 + 0.6 + 0.58 + 0.52 + 0.59}{7} = 0.58$$

Finally, by Corollary 1, we conclude that:

$$0.54 \leq \eta^{c_\varepsilon}(DA_T(S)) \leq 0.6$$

for all $\varepsilon \in [0, 1]$. Hence, due to the certainty of the rules are very similar, the c_ε-efficiency of the algorithm $DA_T(S)$ of this particular example is ranked among two closed values. The most representative value is 0.54, which really measures the classification efficiency of the algorithm $DA_T(S)$, since it corresponds to the average of all the certainties. We can decrease the threshold ε to match similar antecedents (avoiding possible noise) and increase the c_ε-efficiency. □

As a consequence, the introduction of Definition 9 has enabled us to directly interpret the different values of the c_ε-efficiency of a given decision algorithm, which implies a significant advantage with respect to Definition 8. Moreover, Proposition 1 and Corollary 1 have also contributed to analyze the new notion of efficiency. Moreover, this notion will be useful for evaluating different algorithms obtained from the same dataset. This goal will be studied in-depth in the near future.

4 Conclusions and further work

We have introduced a new definition of efficiency of decision algorithms in FRST. This notion is based on the certainties of decision rules and provides the quality of classification of the given algorithm. We have also presented some interesting properties of this new notion, such as the monotony in the threshold and the interval where the efficiency can take values depending on the decision algorithm under consideration.

In the future, we will study the c_ε-efficiency of different decision algorithms to compare them. This fact is very important for a better understanding about the simplification of decision rules, and consequently, the reduction of attributes in the fuzzy framework. Furthermore, we are interested in analyzing real datasets by computing different c_ε-efficiencies to extract significant conclusions.

References

1. M. J. Benítez-Caballero, J. Medina, and E. Ramírez-Poussa. Unifying reducts in formal concept analysis and rough set theory. *Studies in Computational Intelligence*, 796:89–95, 2019.
2. F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Fuzzy rough set decision algorithms. *Communications in Computer and Information Science*, 1601:63–76, 2022.

3. F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Efficiency of fuzzy rough set decision algorithms. *Studies in Computational Intelligence*, 1127:17–24, 2024.
4. F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Rough set decision algorithms for modeling with uncertainty. *Journal of Computational and Applied Mathematics*, 437:115413, 2024.
5. M. D. Cock, C. Cornelis, and E. E. Kerre. Fuzzy rough sets: The forgotten step. *Fuzzy Systems, IEEE Transactions on*, 15(1):121–130, feb. 2007.
6. C. Cornelis, J. Medina, and N. Verbiest. Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection. *International Journal of Approximate Reasoning*, 55:412–426, 2014.
7. A. De Luca and S. Termini. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Information and Control*, 20(4):301–312, 1972.
8. D. Dubois and H. Prade. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17(2-3):191–209, 1990.
9. Y. Gao, D. Chen, and H. Wang. Optimal granularity selection based on algorithm stability with application to attribute reduction in rough set theory. *Information Sciences*, 654:119845, 2024.
10. 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.
11. J. Medina. Multi-adjoint property-oriented and object-oriented concept lattices. *Information Sciences*, 190:95–106, 2012.
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. Information systems theoretical foundations. *Information Systems*, 6(3):205 – 218, 1981.
17. Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.
18. Z. Pawlak. Rough sets and decision algorithms. In W. Ziarko and Y. Yao, editors, *Rough Sets and Current Trends in Computing*, pages 30–45, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.
19. S. Stawicki, D. Ślęzak, A. Janusz, and S. Widz. Decision bireducts and decision reducts - a comparison. *International Journal of Approximate Reasoning*, 84:75–109, 2017.
20. A. Theerens and C. Cornelis. Fuzzy rough sets based on fuzzy quantification. *Fuzzy Sets and Systems*, 473:108704, 2023.
21. C. Y. Wang. Topological structures of L-fuzzy rough sets and similarity sets of l-fuzzy relations. *International Journal of Approximate Reasoning*, 83:160–175, 2017.
22. L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
23. L. Zedam, H. Bouremel, and B. De Baets. Left- and right-compatibility of order relations and fuzzy tolerance relations. *Fuzzy Sets and Systems*, 360:65–81, 2019. Theme: Fuzzy Relations.