axioms

MDPI

Article

A Novel Multi-Criteria Decision-Making Method Based on Rough Sets and Fuzzy Measures

Jingqian Wang [1] and Xiaohong Zhang [1,2,*]

[1] School of Mathematics and Data Science, Shaanxi University of Science and Technology, Xi'an 710021, China; wangjingqianw@163.com
[2] Shaanxi Joint Laboratory of Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China
* Correspondence: zhangxiaohong@sust.edu.cn or zxhonghz@263.net

Abstract: Rough set theory provides a useful tool for data analysis, data mining and decision making. For multi-criteria decision making (MCDM), rough sets are used to obtain decision rules by reducing attributes and objects. However, different reduction methods correspond to different rules, which will influence the decision result. To solve this problem, we propose a novel method for MCDM based on rough sets and a fuzzy measure in this paper. Firstly, a type of non-additive measure of attributes is presented by the importance degree in rough sets, which is a fuzzy measure and called an attribute measure. Secondly, for a decision information system, the notion of the matching degree between two objects is presented under an attribute. Thirdly, based on the notions of the attribute measure and matching degree, a Choquet integral is constructed. Moreover, a novel MCDM method is presented by the Choquet integral. Finally, the presented method is compared with other methods through a numerical example, which is used to illustrate the feasibility and effectiveness of our method.

Keywords: rough set; fuzzy measure; multi-criteria decision making; Choquet integral; attribute reduction

check for updates

Citation: Wang, J.; Zhang, X. A Novel Multi-Criteria Decision-Making Method Based on Rough Sets and Fuzzy Measures. *Axioms* **2022**, *11*, 275. https://doi.org/10.3390/axioms11060275

Academic Editor: Darjan Karabašević

Received: 8 May 2022
Accepted: 1 June 2022
Published: 6 June 2022

1. Introduction

In 1982, Pawlak [1,2] proposed rough set theory, as a mathematical tool, to deal with various kinds of data in data mining. It has been applied in various issues, such as attribute reduction [3–5], rule extraction [6–8], knowledge discovery [9–11] and feature selection [12–14]. To broaden the application ability of Pawlak's rough set theory in practical problems [11,15], it has been extended by generalized relations [16,17], various coverings [18–20] and several types of neighborhoods [4,21]. Moreover, it has been combined with several theories, including lattice theory [22], matrix theory [23], fuzzy set theory [24] and others [25,26].

In multi-criteria decision-making (MCDM) problems [27], it is difficult to obtain the optimal attribute weight. Hence, different attribute weights will influence the decision results. Pawlak's rough sets can obtain decision rules to make decisions, which can solve the issue above. Therefore, the decision-making methods based on Pawlak's rough sets have received more and more attention [28,29]. The decision rules are obtained by reducing attributes and objects in Pawlak's rough sets. Hence, there are many attribute and object reduction methods, such as the discernibility matrix method [30,31], positive region method [32,33], information entropy method [34,35] and other methods [36,37]. Different reduction methods correspond to different rules, which will influence the decision result. Hence, for the existing rule extraction algorithms of rough sets, the decision value will be not unique. For example, we use the hiring dataset taken from Komorowski et al. in [38], where all the attributes have nominal values. We use two famous rule extraction algorithms of rough sets, which are the CN2 algorithm [39] and the LEM2 algorithm [40], to illustrate this statement. We use the R programming language for these two algorithms (the CN2 algorithm [39] and the LEM2 algorithm [40] are at pages 97 and 105 in the package 'RoughSets', respectively). The package 'RoughSets' can be downloaded from https://CRAN.R-project.org/package=RoughSets accessed on 23 May 2022, and the steps of them are shown as follows: firstly, we use the first seven records to obtain rules by the CN2 algorithm [39] and the LEM2 algorithm [40], respectively. Then, we use the obtained rules to make decisions for the eight records x_8. The corresponding results are shown in Section 5.3, and we find that the predicted decision value of x_8 is not unique by using different values in the CN2 algorithm [39] and the LEM2 algorithm [40], respectively. Hence, for a problem of MCDM, different decision rules will influence the decision results. It is necessary to seek a new method of decision making by rough sets.

The research motivations of this paper are listed as follows:

1. In rough set theory, the common decision-making method is using decision rules. It is difficult to find the best decision rules, because different methods can obtain different rules, which will influence the decision result. Hence, a new decision-making method based on rough sets should be presented, which will be independent of decision rules.
2. In decision-making theory, attribute weights are needed in almost all decision-making methods, such as the WA, OWA and TOPSIS methods. However, it is difficult to obtain the optimal weight value, and many weight values are given artificially. To solve this problem, Choquet integrals can be used to aggregate decision information without attribute weights.

In this paper, a novel MCDM method based on rough sets and fuzzy measures is presented. Firstly, to show the correlation between attributes in a decision information system, a type of non-additive measure of attributes is presented by the importance degree in rough sets. It is called an attribute measure, and some properties of it are presented. Secondly, to describe how close any two objects are to each other in a decision information system, the notion of the matching degree between two objects is presented under an attribute. Thirdly, a Choquet integral is constructed based on the notions of attribute measure and matching degree above. Moreover, a novel MCDM method is presented by the Choquet integral, which can aggregate all information between two objects. Finally, to illustrate the feasibility and effectiveness of our method above, our method is compared with other methods through a numerical example. By the corresponding analysis, our method can address the deficiency of the existing methods well.

The rest of this article is organized as follows: Section 2 recalls several basic notions about Pawlak's rough sets, fuzzy measures and Choquet integrals. In Section 3, a type of non-additive measure of attributes is presented by the importance degree in rough sets. Moreover, the notion of the matching degree between two objects is presented under an attribute, as well as corresponding Choquet integrals. In Section 4, a novel MCDM method is presented by the Choquet integral. In Section 5, we show the effectiveness and the efficiency of our method by a numerical example. Section 6 concludes this article and indicates further works.

2. Basic Definitions

In this section, we recall several concepts in Pawlak's rough sets, fuzzy measures and Choquet integrals.

2.1. Pawlak's Rough Sets

We show some notions about Pawlak's rough sets in [1,41] as follows:

Let $S = (U, A)$ be an information system, where U is a nonempty finite set of objects and called the universe, and A is a nonempty finite set of attributes such that $a : U \to V_a$ for any $a \in A$, where V_a is called the value set of a. The indiscernibility relation induced by A is defined as follows:

$$IND(A) = \{(x, y) \in U \times U : \forall a \in A, a(x) = a(y)\}.$$

For every $X \subseteq U$, a pair of approximations $\overline{A}(X)$ and $\underline{A}(X)$ of X are denoted as

$$\overline{A}(X) = \{x \in U : [x]_A \bigcap X \neq \emptyset\},$$
$$\underline{A}(X) = \{x \in U : [x]_A \subseteq X\},$$

where $[x]_A = \{y \in U : (x, y) \in IND(A)\}$ and $U/A = \{[x]_A : x \in U\}$. $\overline{A}$ and $\underline{A}$ are called the upper and lower approximation operators with respect to A, respectively.

Let $\emptyset$ be the empty set and $-X = U - X$. We have the following conclusions about $\overline{A}$ and $\underline{A}$.

Proposition 1 ([1,41])**.** *Let $S = (U, A)$ be an information system. For any $X, Y \subseteq U$,*

(1L) $\underline{A}(U) = U$ *(1H)* $\overline{A}(U) = U$
(2L) $\underline{A}(\phi) = \phi$ *(2H)* $\overline{A}(\phi) = \phi$
(3L) $\underline{A}(X) \subseteq X$ *(3H)* $X \subseteq \overline{A}(X)$
(4L) $\underline{A}(X \bigcap Y) = \underline{A}(X) \bigcap \underline{A}(Y)$ *(4H)* $\overline{A}(X \bigcup Y) = \overline{A}(X) \bigcup \overline{A}(Y)$
(5L) $\underline{A}(\underline{A}(X)) = \underline{A}(X)$ *(5H)* $\overline{A}(\overline{A}(X)) = \overline{A}(X)$
(6L) $X \subseteq Y \Rightarrow \underline{A}(X) \subseteq \underline{A}(Y)$ *(6H)* $X \subseteq Y \Rightarrow \overline{A}(X) \subseteq \overline{A}(Y)$
(7L) $\underline{A}(-\underline{A}(X)) = -\underline{A}(X)$ *(7H)* $\overline{A}(-\overline{A}(X)) = -\overline{A}(X)$
(8LH) $\underline{A}(-X) = -\overline{A}(X)$ *(9LH)* $\underline{A}(X) \subseteq \overline{A}(X)$

Moreover, Let $S = (U, A)$ be an information system. For any $B, C \in A$ and $X \in U$,

$$\underline{B}(X) \bigcup \underline{C}(X) \supseteq \underline{B \bigcup C}(X),$$
$$\underline{B}(X) \bigcap \underline{C}(X) \subseteq \underline{B \bigcap C}(X).$$

Then, $S = (U, A \bigcup D)$ is called a decision information system, where A is a conditional attribute set and D is a decision attribute set. The notions of dependency degree and importance degree in the decision information system are shown in the following definition.

Definition 1 ([1,41])**.** *Let $S = (U, A \bigcup D)$ be a decision information system. Then, the dependency degree of D with regard to A in S is*

$$\gamma_D(A) = \frac{|POS_A(D)|}{|U|} = \frac{\sum_{X \in U/D} |\underline{A}(X)|}{|U|},$$

where $POS_A(D) = \bigcup_{X \in U/D} \underline{A}(X)$. For any $B \in A$, the importance degree of D with regard to B in S is

$$Sig_D(B) = \gamma_D(A) - \gamma_D(A - B).$$

2.2. Fuzzy Measures and Choquet Integrals

Firstly, the definition of the fuzzy measure is shown in Definition 2.

Definition 2 ([42,43])**.** *Given a universe U and a set function $\mathfrak{m}$: $P(U) \to [0, 1]$, where $P(U)$ is the power set of U, $\mathfrak{m}$ is called a fuzzy measure on U if the following statements hold:*

(1) $\mathfrak{m}(\emptyset) = 0$, $\mathfrak{m}(U) = 1$;
(2) $A, B \subseteq U$, $A \subseteq B$, *which implies* $\mathfrak{m}(A) \leq \mathfrak{m}(B)$.

Inspired by the notion of the fuzzy measure, a type of fuzzy integral is proposed in Definition 3.

Definition 3 ([44,45])**.** *Given a real-valued function $f : U \to [0, 1]$ with $U = \{x_1, x_2, \cdots, x_n\}$, the Choquet integral of f with respect to the fuzzy measure $\mathfrak{m}$ is defined as:*

$$\int f d\mathfrak{m} = \sum_{i=1}^{n} [\mathfrak{m}(X_{(i)}) - \mathfrak{m}(X_{(i+1)})] f(x_{(i)}),$$

where $\{x_{(1)}, x_{(2)}, \cdots, x_{(n)}\}$ is a permutation of $\{x_1, x_2, \cdots, x_n\}$ such that $f(x_{(1)}) \leq f(x_{(2)}) \leq \cdots \leq f(x_{(n)})$, $X_{(i)} = \{x_{(i)}, x_{(i+1)}, \cdots, x_{(n)}\}$ and $X_{(n+1)} = \emptyset$.

In Definition 3, the real-valued function $f : U \to [0,1]$ is called a measurable function, which can be seen as a fuzzy set.

3. Fuzzy Rough Measures and Choquet Integrals

In this section, the notions of the attribute measure and matching degree between two objects are presented in a decision information system. The key work of this section is to induce the fuzzy measure and the measurable function from a discrete data table. Based on these new notions, a Choquet integral is constructed.

3.1. Fuzzy Rough Measures Based on Attribute Importance Degrees

In this subsection, a type of non-additive measure of attributes is presented by the importance degree in rough sets, which is a fuzzy measure and called an attribute measure. Moreover, several properties of the attribute measure are proposed. Firstly, the notion of the attribute measure is proposed.

Definition 4. *Let $S = (U, A \bigcup D)$ be a decision information system. For any $B \subseteq A$, we call $\mu(B)$ an attribute measure of B in S, where*

$$\mu(B) = \frac{Sig_D(B)}{Sig_D(A)}.$$

By Definition 4, the notion of the attribute measure reflects the degree of correlation between attribute subset B and attribute set A. It will be a useful tool for describing relational data in rough set theory.

Example 1. *Let $S = (U, A \bigcup D)$ be a decision information system that provides 7 days' meteorological observation data, as shown in Table 1, where A is the set of four attributes of weather, and D denotes whether to hold a meeting. The detailed description of each attribute is as follows:*

- *The conditional attribute '$a_1 = Weatherprediction$' has values: "Clear = 1", "Cloudy = 2", "Rain = 3".*
- *The conditional attribute '$a_2 = Airtemperature$' has values: "Hot = 1", "Warm = 2", "Cool = 3".*
- *The conditional attribute '$a_3 = Windiness$' has values: "Yes = 0", "No = 1".*
- *The conditional attribute '$a_4 = Humidity$' has values: "Wet = 1", "Normal = 2", "Dry = 3".*
- *The conditional attribute 'D' has values: "Yes = 1", "No = 0".*

Table 1. Weather observation data.

U	a_1	a_2	a_3	a_4	D
x_1	1	1	1	1	1
x_2	2	2	1	2	1
x_3	2	2	1	1	1
x_4	1	2	0	3	1
x_5	1	3	1	2	0
x_6	3	3	1	3	0
x_7	2	1	1	2	0

Then, $U/D = \{X_1, X_2\}$, where $X_1 = \{x_1, x_2, x_3, x_4\}$, $X_2 = \{x_5, x_6, x_7\}$. Hence, $\underline{A}(X_1) = \{x_1, x_2, x_3, x_4\}$ and $\underline{A}(X_2) = \{x_5, x_6, x_7\}$.

By Definition 1,

$$POS_A(D) = \underline{A}(X_1) \bigcup \underline{A}(X_2)) = U, \text{ i.e., } \gamma_D(A) = \tfrac{|POS_A(D)|}{|U|} = 1.$$

Thus,

$$Sig_D(A) = \gamma_D(A) - \gamma_D(\varnothing) = 1 - 0 = 1.$$

Suppose $B = \{a_1, a_2\}$. We have

$$\underline{A-B}(X_1) = \{x_1, x_3, x_4\} \text{ and } \underline{A-B}(X_2) = \{x_6\}.$$

By Definition 1,

$$POS_{A-B}(D) = \underline{A-B}(X_1) \bigcup \underline{A-B}(X_2) = \{x_1, x_3, x_4, x_6\}, \text{ i.e.,}$$
$$\gamma_D(A) = \frac{|POS_{A-B}(D)|}{|U|} = \frac{4}{7} = 0.5714.$$

Hence,

$$Sig_D(B) = \gamma_D(A) - \gamma_D(A-B) = 1 - 0.5714 = 0.4286.$$

Therefore, by Definition 4, we have

$$\mu(B) = \frac{Sig_D(B)}{Sig_D(A)} = \frac{0.4286}{1} = 0.4286.$$

Several properties of the attribute measure in Definition 4 are proposed below.

Proposition 2. *Let $S = (U, A \bigcup D)$ be a decision information system, and $\mu(B)$ be a attribute measure for any $B \subseteq A$. Then,*

(1) $\mu(\emptyset) = 0$ *and* $\mu(A) = 1$;
(2) *For any $B, C \subseteq A$, $B \subseteq C$ implies $\mu(B) \leq \mu(C)$.*

Proof. (1) By Definition 1 and Proposition 1, we have that $\gamma_D(\emptyset) = 0$ and $\gamma_D(A) \neq 0$. Hence,

$$\mu(\emptyset) = \frac{Sig_D(\emptyset)}{Sig_D(A)} = \frac{\gamma_D(A) - \gamma_D(A)}{\gamma_D(A) - \gamma_D(\emptyset)} = 0, \mu(A) = \frac{Sig_D(A)}{Sig_D(A)} = \frac{\gamma_D(A) - \gamma_D(\emptyset)}{\gamma_D(A) - \gamma_D(\emptyset)} = 1.$$

(2) For any $B, C \subseteq A$ and $X \in U$, if $B \subseteq C$, then $\underline{A-C}(X) \subseteq \underline{A-B}(X)$ by Proposition 1. Hence, $\gamma_D(A-C) \leq \gamma_D(A-B)$, i.e., $Sig_D(B) \leq Sig_D(C)$. Therefore,

$$\mu(B) = \frac{Sig_D(B)}{Sig_D(A)} \leq \frac{Sig_D(C)}{Sig_D(A)} = \mu(C), \text{ i.e., } \mu(B) \leq \mu(C).$$

□

Example 2 (Continued from Example 1). *Let $C = \{a_1, a_2, a_3\}$. $\underline{A-C}(X_1) = \{x_1, x_3\}$ and $\underline{A-C}(X_2) = \emptyset$. By Definition 1, $POS_{A-C}(D) = \underline{A-C}(X_1) \bigcup \underline{A-C}(X_2) = \{x_1, x_3\}$, i.e., $\gamma_D(C) = \frac{|POS_{A-C}(D)|}{|U|} = \frac{2}{7} = 0.2857$. $Sig_D(C) = \gamma_D(A) - \gamma_D(A-C) = 1 - 0.2857 = 0.7143$. Hence,*

$$\mu(C) = \frac{Sig_D(C)}{Sig_D(A)} = 0.7143/1 = 0.7143.$$

Therefore, $B \subseteq C$ implies $\mu(B) \leq \mu(C)$.

Proposition 3. *Let $S = (U, A \bigcup D)$ be a decision information system, and $\mu(B)$ be a attribute measure for any $B \subseteq A$. Then, $0 \leq \mu(B) \leq 1$.*

Proof. By Proposition 1 and the statement (2) in Proposition 2, $\mu(\emptyset) \leq \mu(B) \leq \mu(A)$. According to (1) in Proposition 2, $0 \leq \mu(B) \leq 1$. □

Example 3 (Continued from Example 1). *In Examples 1 and 2, $\mu(B) = 0.4286$ and $\mu(C) = 0.7143$. Hence, $0 \leq \mu(B), \mu(C) \leq 1$.*

Proposition 4. *Let $S = (U, A \bigcup D)$ be a decision information system, and $\mu(B)$ and $\mu(C)$ be two attribute measures for any $B, C \subseteq A$. Then, $\mu(B) + \mu(C) \geq 2\mu(B \bigcap C)$.*

Proof. By the statement (2) in Proposition 2, $\mu(B) \geq \mu(B \bigcap C)$ and $\mu(C) \geq \mu(B \bigcap C)$. Hence, $\mu(B) + \mu(C) \geq 2\mu(B \bigcap C)$. □

Example 4 (Continued from Example 1). *In Examples 1 and 2, $\mu(B) = 0.4286$ and $\mu(C) = 0.7143$. Since $\mu(B \bigcap C) = 0.4286$, $\mu(B) + \mu(C) \geq 2\mu(B \bigcap C)$.*

Proposition 5. *Let $S = (U, A \bigcup D)$ be a decision information system, and $\mu(B)$ and $\mu(C)$ be two attribute measures for any $B, C \subseteq A$. Then, $\mu(B) + \mu(C) \leq 2\mu(A \bigcup B)$.*

Proof. By the statement (2) in Proposition 2, $\mu(B) \leq \mu(B \bigcup C)$ and $\mu(C) \leq \mu(B \bigcup C)$. Hence, $\mu(B) + \mu(C) \leq 2\mu(B \bigcup C)$. □

Example 5 (Continued from Example 1). *In Examples 1 and 2, $\mu(B) = 0.4286$ and $\mu(C) = 0.7143$. Since $\mu(B \bigcup C) = 0.7143$, $\mu(B) + \mu(C) \leq 2\mu(B \bigcup C)$.*

Theorem 1. *Let $S = (U, A \bigcup D)$ be a decision information system, and $\mu(B)$ be a attribute measure for any $B \subseteq A$. Then, μ is a fuzzy measure on A.*

Proof. By Proposition 3, we find that μ is a set function where $\mu : P(A) \to [0, 1]$. According to Proposition 2, the statements (1) and (2) in Definition 4 hold for μ. Hence, μ is a fuzzy measure on A. □

Inspired by Theorem 1, we also call μ a fuzzy rough measure in a decision information system $S = (U, A \bigcup D)$. In Example 5, we find that $\mu(B) + \mu(C) \neq \mu(B \bigcup C)$. Hence, μ is a non-additive measure, which shows that attributes are related in the decision information system $S = (U, A \bigcup D)$.

3.2. Choquet Integrals under Fuzzy Rough Measures

In this subsection, for a decision information system, the notion of the matching degree between two objects is presented under an attribute. Based on the notions of attribute measure and matching degree, a Choquet integral is constructed.

Definition 5. *Let $S = (U, A \bigcup D)$ be a decision information system. For any $x, y \in U$ and $a \in A$, we call $f_{(x,y)}(a)$ the matching degree between x and y with respect to a, where*

$$f_{(x,y)}(a) = \frac{1}{1 + |a(x) - a(y)|}.$$

Example 6 (Continued from Example 1). *By Definition 5, we have*

$$\begin{aligned} f_{(x_1,x_2)}(a_1) &= \frac{1}{1 + |a_1(x_1) - a_1(x_2)|} = 0.5. \\ f_{(x_1,x_2)}(a_2) &= \frac{1}{1 + |a_2(x_1) - a_2(x_2)|} = 0.5. \\ f_{(x_1,x_2)}(a_3) &= \frac{1}{1 + |a_3(x_1) - a_3(x_2)|} = 1.0. \\ f_{(x_1,x_2)}(a_4) &= \frac{1}{1 + |a_4(x_1) - a_4(x_2)|} = 0.5. \end{aligned}$$

Theorem 2. *Let $S = (U, A \bigcup D)$ be a decision information system with $A = \{a_1, a_2, \cdots, a_n\}$, and μ be a fuzzy rough measure in $S = (U, A \bigcup D)$. Then, for any $x, y \in U$,*

$$\int f_{(x,y)} d\mu = \sum_{i=1}^{n} [\mu(A_{(i)}) - \mu(A_{(i+1)})] f_{(x,y)}(a_{(i)}),$$

is a Choquet integral of $f_{(x,y)}$ with respect to the fuzzy rough measure μ on A, where $\{a_{(1)}, a_{(2)}, \cdots, a_{(n)}\}$ is a permutation of $\{a_1, a_2, \cdots, a_n\}$ such that $f_{(x,y)}(a_{(1)}) \leq f_{(x,y)}(a_{(2)}) \leq \cdots \leq f_{(x,y)}(a_{(n)})$, $A_{(i)} = \{a_{(i)}, a_{(i+1)}, \cdots, a_{(n)}\}$ and $A_{(n+1)} = \emptyset$.

Proof. By Theorem 1, we know that the fuzzy rough measure μ is a fuzzy measure. Hence, it is immediate by Definition 3. □

Remark 1. *In Theorem 2, we find that $\mu(A_{(i)})$, $\mu(A_{(i+1)})$ and $a_{(i)}$ are related to $f_{(x,y)}$. Therefore, we denote $\mu(A_{(i)})$, $\mu(A_{(i+1)})$ and $a_{(i)}$ by $\mu(A_{(i)}^{f_{(x,y)}})$, $\mu(A_{(i+1)}^{f_{(x,y)}})$ and $a_{(i)}^{f_{(x,y)}}$ in the following discussion.*

Example 7 (Continued from Example 1). *By Example 6, we have*

$$f_{(x_1,x_2)}(a_1) \le f_{(x_1,x_2)}(a_2) \le f_{(x_1,x_2)}(a_4) \le f_{(x_1,x_2)}(a_3).$$

Hence, for $f_{(x_1,x_2)}$, we obtain

$$a_{(1)}^{f_{(x_1,x_2)}} = a_1,\ a_{(2)}^{f_{(x_1,x_2)}} = a_2,\ a_{(3)}^{f_{(x_1,x_2)}} = a_4 \text{ and } a_{(4)}^{f_{(x_1,x_2)}} = a_3.$$

Hence,

$$A_{(i)}^{f_{(x_1,x_2)}} = \{a_{(i)}^{f_{(x_1,x_2)}}, a_{(i+1)}^{f_{(x_1,x_2)}}, \cdots, a_{(4)}^{f_{(x_1,x_2)}}\}\ (i = 1,2,3,4) \text{ and } A_{(5)}^{f_{(x_1,x_2)}} = \emptyset.$$

Therefore, by Definition 4, we have

$$\mu(A_{(1)}^{f_{(x_1,x_2)}}) = 1,\ \mu(A_{(2)}^{f_{(x_1,x_2)}}) = 0.8571,\ \mu(A_{(3)}^{f_{(x_1,x_2)}}) = 0,\ \mu(A_{(4)}^{f_{(x_1,x_2)}}) = 0 \text{ and } \mu(A_{(5)}^{f_{(x_1,x_2)}}) = 0.$$

By Theorem 2,

$$\int f_{(x_1,x_2)} d\mu = (1 - 0.8571) \times 0.5 + (0.8571 - 0) \times 0.5 + 0 \times 0.5 + 0 \times 1 = 0.5.$$

In the same way, we have

$$\begin{aligned}
\int f_{(x_2,x_2)} d\mu &= \sum_{i=1}^{4} [\mu(A_{(i)}^{f_{(x_2,x_2)}}) - \mu(A_{(i+1)}^{f_{(x_2,x_2)}})] f_{(x_2,x_2)}(a_{(i)}^{f_{(x_2,x_2)}}) = 1.0, \\
\int f_{(x_3,x_2)} d\mu &= \sum_{i=1}^{4} [\mu(A_{(i)}^{f_{(x_3,x_2)}}) - \mu(A_{(i+1)}^{f_{(x_3,x_2)}})] f_{(x_3,x_2)}(a_{(i)}^{f_{(x_3,x_2)}}) = 0.8571, \\
\int f_{(x_4,x_2)} d\mu &= \sum_{i=1}^{4} [\mu(A_{(i)}^{f_{(x_4,x_2)}}) - \mu(A_{(i+1)}^{f_{(x_4,x_2)}})] f_{(x_4,x_2)}(a_{(i)}^{f_{(x_4,x_2)}}) = 0.6429, \\
\int f_{(x_5,x_2)} d\mu &= \sum_{i=1}^{4} [\mu(A_{(i)}^{f_{(x_5,x_2)}}) - \mu(A_{(i+1)}^{f_{(x_5,x_2)}})] f_{(x_5,x_2)}(a_{(i)}^{f_{(x_5,x_2)}}) = 0.5, \\
\int f_{(x_6,x_2)} d\mu &= \sum_{i=1}^{4} [\mu(A_{(i)}^{f_{(x_6,x_2)}}) - \mu(A_{(i+1)}^{f_{(x_6,x_2)}})] f_{(x_6,x_2)}(a_{(i)}^{f_{(x_6,x_2)}}) = 0.5, \\
\int f_{(x_7,x_2)} d\mu &= \sum_{i=1}^{4} [\mu(A_{(i)}^{f_{(x_7,x_2)}}) - \mu(A_{(i+1)}^{f_{(x_7,x_2)}})] f_{(x_7,x_2)}(a_{(i)}^{f_{(x_7,x_2)}}) = 0.6429.
\end{aligned}$$

In Example 7, we have that $\int f_{(x_2,x_2)} d\mu = 1.0$, which is greater than other values of $\int f_{(x_j,x_2)} d\mu$ $(j = 1,3,4,5,6,7)$. $\int f_{(x_2,x_2)} d\mu = 1.0$ means that x_2 is the best match to itself, which is consistent with actual logic.

4. A Novel Decision-Making Method Based on Fuzzy Rough Measures and Choquet Integrals

In this section, a novel MCDM method is presented by the Choquet integral, which can aggregate all information between two objects.

4.1. The Problem of Decision Making

Let $S = (U, A \bigcup D)$ be a decision information system, which is shown in Table 2, where $U = \{x_1, \cdots, x_m\}$ is the set of objects, $A = \{a_1, \cdots, a_n\}$ is a conditional attribute set, D is a decision attribute, $x_{ji} = a_i(x_j)$ is the attribute value of x_j under conditional attribute a_j, and d_j is the decision value of x_j under decision attribute D. For a new object x_{m+1}, we take the value of each conditional attribute to be $(a_1(x_{m+1}))$, $(a_2(x_{m+1}))$, $\cdots$, $(a_n(x_{m+1}))$. Then, the decision maker should give the decision value of x_{m+1} according to $S = (U, A \bigcup D)$.

Table 2. A decision-making table.

U	a_1	a_2	$\cdots$	a_n	D
x_1	x_{11}	x_{12}	$\cdots$	x_{1n}	d_1
x_2	x_{21}	x_{22}	$\cdots$	x_{2n}	d_2
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
x_m	x_{m1}	x_{m2}	$\cdots$	x_{mn}	d_m

4.2. The Novel Decision-Making Method

Based on Theorems 1 and 2, we present a novel method to solve the issue of MCDM by using fuzzy rough measures and Choquet integrals. We show this novel method as follows, for the problem of decision making in Section 4.1:

Step 1: For any $x_j \in U$ $(j = 1, 2, \cdots, m)$ and $a_i \in A$ $(i = 1, 2, \cdots, n)$, we calculate all matching degrees $f_{(x_j, x_{m+1})}(a_i) = \dfrac{1}{1 + |a_j(x_j) - a_i(x_{m+1})|}$, which are shown in Table 3.

Table 3. A matching degree table.

U	a_1	a_2	$\cdots$	a_n
x_1	$f_{(x_1,x_{m+1})}(a_1)$	$f_{(x_1,x_{m+1})}(a_2)$	$\cdots$	$f_{(x_1,x_{m+1})}(a_n)$
x_2	$f_{(x_2,x_{m+1})}(a_1)$	$f_{(x_2,x_{m+1})}(a_2)$	$\cdots$	$f_{(x_2,x_{m+1})}(a_n)$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
x_m	$f_{(x_m,x_{m+1})}(a_1)$	$f_{(x_m,x_{m+1})}(a_2)$	$\cdots$	$f_{(x_m,x_{m+1})}(a_n)$

Step 2: For any $x_j \in U$ $(j = 1, 2, \cdots, m)$, we calculate all Choquet integrals under fuzzy rough measures μ, which are shown as follows:

$$s(x_1, x_{m+1}) = \int f_{(x_1,x_{m+1})} d\mu = \sum_{i=1}^{n} [\mu(A_{(i)}^{f_{(x_1,x_{m+1})}}) - \mu(A_{(i+1)}^{f_{(x_1,x_{m+1})}})] f_{(x_1,x_{m+1})}(a_{(i)}^{f_{(x_1,x_{m+1})}}),$$

$$\vdots$$

$$s(x_j, x_{m+1}) = \int f_{(x_j,x_{m+1})} d\mu = \sum_{i=1}^{n} [\mu(A_{(i)}^{f_{(x_j,x_{m+1})}}) - \mu(A_{(i+1)}^{f_{(x_j,x_{m+1})}})] f_{(x_j,x_{m+1})}(a_{(i)}^{f_{(x_j,x_{m+1})}}),$$

$$\vdots$$

$$s(x_m, x_{m+1}) = \int f_{(x_m,x_{m+1})} d\mu = \sum_{i=1}^{n} [\mu(A_{(i)}^{f_{(x_m,x_{m+1})}}) - \mu(A_{(i+1)}^{f_{(x_m,x_{m+1})}})] f_{(x_m,x_{m+1})}(a_{(i)}^{f_{(x_m,x_{m+1})}}).$$

Step 3: We obtain the ranking of all alternatives by the value of $s(x_j, x_{m+1})$. Moreover, the decision maker chooses the best one whose decision value is the same as that of x_{m+1}.

For steps 1–3 above, the MCDM algorithm by fuzzy rough measures and Choquet integrals is shown in Algorithm 1.

Algorithm 1 The MCDM algorithm by fuzzy rough measures and Choquet integrals

Input: A decision information system $S = (U, A \bigcup D)$ and a new decision object x_{m+1}, where $U = \{x_1, \cdots, x_m\}$, $A = \{a_1, \cdots, a_n\}$.
Output: The decision value of x_{m+1}.
(1) for $j = 1 \to m$
(2) for $i = 1 \to m$
(3) Compute $f_{(x_j, x_{m+1})}(a_i)$;
(4) end
(5) Compute $s(x_j, x_{m+1}) = \int f_{(x_j, x_{m+1})} d\mu$;
(6) end
(7) for $j = 1 \to m$
(8) Obtain the ranking of all $s(x_j, x_{m+1})$;
(9) end
(10) Give the decision value of x_{m+1} by the ranking of all $s(x_j, x_{m+1})$.

5. Comparison and Analysis

To illustrate the feasibility and effectiveness of our method above, it is compared with other methods through a numerical example in this section.

5.1. Hiring Dataset

In this section, we list the hiring dataset taken from Komorowski et al. in [38], where all the attributes have nominal values, which is shown in Table 4. It contains 8 objects with 4 conditional attributes and 1 decision attribute. The detailed description of each attribute is as follows:

- The conditional attribute 'Diploma' has values: "MBA", "MSc", "MCE".
- The conditional attribute 'Experience' has values: "High", "Low", "Medium".
- The conditional attribute 'French' has values: "Yes", "No".
- The conditional attribute 'Reference' has values: "Excellent", "Good", "Neutral".
- The conditional attribute 'Decision' has values: "Accept", "Reject".

Table 4. The hiring dataset [38].

U	Diploma	Experience	French	Reference	Decision
x_1	MBA	Medium	Yes	Excellent	Accept
x_2	MSC	High	Yes	Neutral	Accept
x_3	MSC	High	Yes	Excellent	Accept
x_4	MBA	High	No	Good	Accept
x_5	MBA	Low	Yes	Neutral	Reject
x_6	MCE	Low	Yes	Good	Reject
x_7	MSC	Medium	Yes	Neutral	Reject
x_8	MCE	Low	No	Excellent	Reject

5.2. An Applied Example

For the hiring dataset [38], which is shown in Table 4 in the paper, we denote the first seven records as the original decision information, and the eighth record x_8 as a new object (we suppose that we do not know the decision value of x_8). In order to facilitate the calculation, we perform the following for Table 4:

- The conditional attribute 'Diploma = a_1' has values: "MBA = 1", "MSc = 2", "MCE = 3".
- The conditional attribute 'Experience = a_2' has values: "Medium = 1", "High = 2", "Low = 3".
- The conditional attribute 'French = a_3' has values: "Yes = 1", "No = 0".
- The conditional attribute 'Reference = a_4' has values: "Excellent = 1", "Neutral = 2", "Good = 3".
- The conditional attribute 'Decision = D' has values: "Accept = 1", "Reject = 0".

It can be denoted as in Table 5.

Table 5. A decision problem in the hiring dataset.

	U	Diploma (a_1)	Experience (a_2)	French (a_3)	Reference (a_4)	Decision (D)
An information system $IS = (U', A)$	x_1	MBA (1)	Medium (1)	Yes (1)	Excellent (1)	Accept (1)
	x_2	MSC (2)	High (2)	Yes (1)	Neutral (2)	Accept (1)
	x_3	MSC (2)	High (2)	Yes (1)	Excellent (1)	Accept (1)
	x_4	MBA (1)	High (2)	No (0)	Good (3)	Accept (1)
	x_5	MBA (1)	Low (3)	Yes (1)	Neutral (2)	Reject (0)
	x_6	MCE (3)	Low (3)	Yes (1)	Good (3)	Reject (0)
	x_7	MSC (2)	Medium (1)	Yes (1)	Neutral (2)	Reject (0)
A decision object	x_8	MCE (3)	Low (3)	No (0)	Excellent (1)	"?"

Then, we use our method to predict the decision value of x_8, i.e., we should predict the "?" in Table 5.

Example 8. *Let $IS = (U, A)$ be an information system, which is the first seven records shown in the hiring dataset [38]. For a decision object x_8, we take the value of each conditional attribute to be $a_1(x_8) = 3$, $a_2(x_8) = 3$, $a_3(x_8) = 0$, $a_4(x_8) = 1$. It is shown in Table 5. Then, we use the following steps to give the decision value of x_8 according to $S = (U, A \bigcup D)$.*

Step 1: *For any $x_j \in U$ ($j = 1, 2, \cdots, 7$) and $a_i \in A$ ($i = 1, 2, 3, 4$), we calculate all matching degrees $f_{(x_j, x_8)}(a_i)$, which are shown in Table 6.*

Table 6. $f_{(x_j,x_8)}(a_i)$, $(i = 1, 2, 3, 4)$ and $(j = 1, 2, \cdots, 7)$.

U	a_1	a_2	a_3	a_4
$f_{(x_1,x_8)}(a_i)$	0.3333	0.3333	0.5000	1.0000
$f_{(x_2,x_8)}(a_i)$	0.5000	0.5000	0.5000	0.5000
$f_{(x_3,x_8)}(a_i)$	0.5000	0.5000	0.5000	1.0000
$f_{(x_4,x_8)}(a_i)$	0.3333	0.5000	1.0000	0.3333
$f_{(x_5,x_8)}(a_i)$	0.3333	1.0000	0.5000	0.5000
$f_{(x_6,x_8)}(a_i)$	1.0000	1.0000	0.5000	0.3333
$f_{(x_7,x_8)}(a_i)$	0.5000	0.3333	0.5000	0.5000

Step 2: *For $f_{(x_1,x_8)}$, we have*

$$f_{(x_1,x_8)}(a_1) \leq f_{(x_1,x_8)}(a_2) \leq f_{(x_1,x_8)}(a_3) \leq f_{(x_1,x_8)}(a_4).$$

Hence, we obtain

$$a_{(1)}^{f_{(x_1,x_8)}} = a_1,\ a_{(2)}^{f_{(x_1,x_8)}} = a_2,\ a_{(3)}^{f_{(x_1,x_8)}} = a_3 \text{ and } a_{(4)}^{f_{(x_1,x_8)}} = a_4.$$

Hence,

$$A_{(i)}^{f_{(x_1,x_8)}} = \{a_{(i)}^{f_{(x_1,x_8)}}, a_{(i+1)}^{f_{(x_1,x_8)}}, \cdots, a_{(4)}^{f_{(x_1,x_8)}}\}\ (i = 1, 2, 3, 4) \text{ and } A_{(5)}^{f_{(x_1,x_8)}} = \emptyset.$$

Therefore, by Definition 4, we have

$$\mu(A_{(1)}^{f_{(x_1,x_8)}}) = 1,\ \mu(A_{(2)}^{f_{(x_1,x_8)}}) = 0.8571,\ \mu(A_{(3)}^{f_{(x_1,x_8)}}) = 0,\ \mu(A_{(4)}^{f_{(x_1,x_8)}}) = 0 \text{ and } \mu(A_{(5)}^{f_{(x_1,x_8)}}) = 0.$$

In the same way, for any $f_{(x_j,x_8)}$ ($j = 1, 2, \cdots, 7$), we can obtain all permutations of $\{a_1, a_2, a_3, a_4\}$ in Table 7.

Table 7. $\{a_{(1)}^{f_{(x_j,x_8)}}, a_{(2)}^{f_{(x_j,x_8)}}, a_{(3)}^{f_{(x_j,x_8)}}, a_{(4)}^{f_{(x_j,x_8)}}\}$ relates to any $f_{(x_j,x_8)}$ ($j = 1, 2, \cdots, 7$).

	$a_{(1)}^{f_{(x_j,x_8)}}$	$a_{(2)}^{f_{(x_j,x_8)}}$	$a_{(3)}^{f_{(x_j,x_8)}}$	$a_{(4)}^{f_{(x_j,x_8)}}$
$f_{(x_1,x_8)}$	a_1	a_2	a_3	a_4
$f_{(x_2,x_8)}$	a_1	a_2	a_3	a_4
$f_{(x_3,x_8)}$	a_1	a_2	a_3	a_4
$f_{(x_4,x_8)}$	a_1	a_4	a_2	a_3
$f_{(x_5,x_8)}$	a_1	a_3	a_4	a_2
$f_{(x_6,x_8)}$	a_4	a_3	a_1	a_2
$f_{(x_7,x_8)}$	a_2	a_1	a_3	a_4

By Table 7, we can calculate all $\mu(A_{(i)}^{f_{(x_j,x_8)}})$ in Table 8, where $i \in \{1, 2, 3, 4\}$, $j = 1, 2, \cdots, 7$ and $\mu(A_{(5)}^{f_{(x_j,x_8)}}) = 0$.

Table 8. $\mu(A_{(i)}^{f_{(x_j,x_8)}})$ with $i \in \{1, 2, 3, 4\}$ and $j = 1, 2, \cdots, 7$.

	$A_{(1)}^{f_{(x_j,x_8)}}$	$A_{(2)}^{f_{(x_j,x_8)}}$	$A_{(3)}^{f_{(x_j,x_8)}}$	$A_{(4)}^{f_{(x_j,x_8)}}$
$\mu(A_{(i)}^{f_{(x_1,x_8)}})$	1	0.8571	0	0
$\mu(A_{(i)}^{f_{(x_2,x_8)}})$	1	0.8571	0	0
$\mu(A_{(i)}^{f_{(x_3,x_8)}})$	1	0.8571	0	0
$\mu(A_{(i)}^{f_{(x_4,x_8)}})$	1	0.8571	0.2857	0
$\mu(A_{(i)}^{f_{(x_5,x_8)}})$	1	0.8571	0.7143	0.2857
$\mu(A_{(i)}^{f_{(x_6,x_8)}})$	1	0.7143	0.4286	0
$\mu(A_{(i)}^{f_{(x_7,x_8)}})$	1	0.2857	0	0

By Table 8 and Theorem 2, we calculate

$$\begin{aligned}
s(x_1, x_8) &= \int f_{(x_1,x_8)} d\mu \\
&= (1 - 0.8571) \times 0.3333 + (0.8571 - 0) \times 0.3333 + 0 \times 0.5 + 0 \times 1 \\
&= 0.3333; \\
s(x_2, x_8) &= \int f_{(x_2,x_8)} d\mu \\
&= (1 - 0.8571) \times 0.5 + (0.8571 - 0) \times 0.5 + 0 \times 0.5 + 0 \times 0.5 \\
&= 0.5000; \\
s(x_3, x_8) &= \int f_{(x_3,x_8)} d\mu \\
&= (1 - 0.8571) \times 0.5 + (0.8571 - 0) \times 0.5 + 0 \times 0.5 + 0 \times 1 \\
&= 0.5000; \\
s(x_4, x_8) &= \int f_{(x_4,x_8)} d\mu \\
&= (1 - 0.8571) \times 0.3333 + (0.8571 - 0.2857) \times 0.3333 + (0.2857 - 0) \times 0.5 + 0 \times 1 \\
&= 0.3810; \\
s(x_5, x_8) &= \int f_{(x_5,x_8)} d\mu \\
&= (1 - 0.8571) \times 0.3333 + (0.8571 - 0.7143) \times 0.5 + (0.7143 - 0.2857) \times 0.5 + (0.2857 - 0) \times 1 \\
&= 0.6190; \\
s(x_6, x_8) &= \int f_{(x_6,x_8)} d\mu \\
&= (1 - 0.7143) \times 0.3333 + (0.7143 - 0.4286) \times 0.5 + (0.4286 - 0) \times 1 + 0 \times 1 \\
&= 0.6667; \\
s(x_7, x_8) &= \int f_{(x_7,x_8)} d\mu \\
&= (1 - 0.2857) \times 0.3333 + (0.0.2857 - 0) \times 0.5 + 0 \times 0.5 + 0 \times 0.5 \\
&= 0.3810.
\end{aligned}$$

Step 3: *We obtain the ranking of all alternatives by the value of* $s(x_j, x_{m+1})$*, where* $s(x_6, x_8)$ *is the best one. Hence, the decision value of* x_8 *is the same as that of* x_6*, which is 0.*

5.3. Comparison with Other Methods

We use the R programming language for dealing with Example 8 by the AQ algorithm [46], the CN2 algorithm [39] and the LEM2 algorithm [40], respectively. The AQ algorithm [46], the CN2 algorithm [39]) and the LEM2 algorithm [40] are at pages 96, 97 and 105 in the the package 'RoughSets', respectively. The package 'RoughSets' can be downloaded from https://CRAN.R-project.org/package=RoughSets, accessed on 23 May 2022. In fact, Table 1 and x_8 are taken from the hiring dataset in [38], where the actual decision value of x_8 is 0. Then, we use some existing algorithms to predict the decision value of x_8 according to Table 1. All results are shown in Table 9.

Table 9. The decision results of x_8 utilizing different methods for Example 8.

Methods	The Actual Decision Value of x_8	The Predicted Decision Value of x_8
The AQ algorithm [46]	0	1
The CN2 algorithm [39]	0	0
The LEM2 algorithm [40]	0	0
Algorithm 1 in this paper [40]	0	0

As shown in Table 9, we find that our method is effective, since the predicted value is equal to the actual value. In the AQ algorithm [46], we use "nOFItervales = 3", "confidence = 0.8" and "timescovered = 3", and then we obtain 6 rules to make a decision. In the CN2 algorithm [39], we use "nOFItervales = 3", and then we obtain two rules to make a decision. In the LEM2 algorithm [40], we use "maxNOfCuts = 1", and then we obtain two rules to make a decision. The AQ algorithm [46], the CN2 algorithm [39] and the LEM2 algorithm [40] all depend on the corresponding rules, which are obtained through rough sets. Although the CN2 algorithm [39] and the LEM2 algorithm [40] can also obtain the predicted value 0 for x_8, the predicted value will be changed by different threshold values. We present some discussions on this statement.

For the AQ algorithm [46], the predicted value is also 1, which does not equal the actual value, although we changed "nOFItervales", "confidence" and "timescovered". For example, we use "nOFItervales = 3", "confidence = 0.9" and "timescovered = 8", and we obtain 16 rules and the predicted value 1; we use "nOFItervales = 1", "confidence = 0.9" and "timescovered = 3", and we obtain 15 rules and the predicted value 1; we use "nOFItervales = 3", "confidence = 0.98" and "timescovered = 28", and we obtain 56 rules and the predicted value 1. Hence, we only present the CN2 algorithm [39] and the LEM2 algorithm [40] in Table 10.

Table 10. The decision results of x_8 utilizing different threshold values for Example 8.

Different Threshold Values in Algorithms	Rules	The Predicted Decision Value Of x_8
"nOFItervales = 3" in the CN2 algorithm [39]	2	0
"nOFItervales = 1" in the CN2 algorithm [39]	6	1
"maxNOfCuts = 1" in the LEM2 algorithm [40]	2	0
"maxNOfCuts = 3" in the LEM2 algorithm [40]	3	1

As shown in Table 10, we find that the predicted decision value of x_8 is changed by using different values in the CN2 algorithm [39] and the LEM2 algorithm [40], respectively. However, our method uses the matching degree between any original object $x_j \in U$ $(j = 1, 2, \cdots, 7)$ and the decision object x_8, and then corresponding Choquet integrals are used to aggregate them. Hence, the result of our method is unique. In particular, our method is more stable than others. For the above comparative analysis, our method is more feasible and effective than others under the hiring dataset [38].

6. Conclusions

In this article, we combine rough sets and fuzzy measures to solve the problem of MCDM, which can well avoid the limitations of the existing decision-making method under rough sets. The contributions of this paper are listed as follows:

- The notion of the attribute measure is presented based on the importance degree in rough sets, which can illustrate the non-additive relationship of two attributes in rough sets. By the new notion, we can find that attributes are related to each other in information systems. It can also be used to construct the corresponding Choquet integral.
- Then, a type of nonlinear aggregation operator (i.e., Choquet integral) is constructed, which can aggregate all information between two objects in a decision information system. Moreover, a method based on the Choquet integral is proposed to deal with the problem of MCDM, which is inspired by case-based reasoning theory. This novel method can address the deficiency of the existing methods well. It can solve the issue of attribute association in MCDM.

In further research, the following topics can be considered: other integrals and generalized rough set models [47–49] will be connected with the research content of this article. The novel method can be combined with other decision-making and aggregation methods [50–52].

Author Contributions: This paper was written with the contribution of all authors. The individual contributions and responsibilities of all authors can be described as follows: J.W. analyzed the existing work on rough sets and fuzzy measures and wrote the paper. X.Z. put forward the idea of this paper and also completed the preparatory work of the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant No. 61976130, and the Natural Science Foundation of Education Department of Shaanxi Province under Grant No. 20JK0506.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: In this paper, we use the hiring dataset taken from Komorowski et al. in [38]. It is also at page 106 in the the package 'RoughSets', respectively. The package 'RoughSets' can be downloaded from https://CRAN.R-project.org/package=RoughSets accessed on 23 May 2022.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pawlak, Z. *Rough Sets: Theoretical Aspects of Reasoning About Data*; Kluwer Academic Publishers: Boston, MA, USA, 1991.
2. Pawlak, Z. Rough sets. *Int. J. Comput. Inf. Sci.* **1982**, *11*, 341–356. [CrossRef]
3. Chen, J.; Lin, Y.; Lin, G.; Li, J.; Ma, Z. The relationship between attribute reducts in rough sets and minimal vertex covers of graphs. *Inf. Sci.* **2015**, *325*, 87–97. [CrossRef]
4. Dai, J.; Hu, Q.; Hu, H. Neighbor inconsistent pair selection for attribute reduction by rough set approach. *IEEE Trans. Fuzzy Syst.* **2018**, *26*, 937–950. [CrossRef]
5. Yang, Y.; Chen, D.; Wang, H. Active sample selection based incremental algorithm for attribute reduction with rough sets. *IEEE Trans. Fuzzy Syst.* **2017**, *25*, 825–838. [CrossRef]
6. Guo, Y.; Tsang, E.C.; Hu, M.; Lin, X.; Chen, D.; Xu, W.; Sang, B. Incremental updating approximations for double-quantitative decision-theoretic rough sets with the variation of objects. *Knowl.-Based Syst.* **2019**, *189*, 105082. [CrossRef]
7. Du, Y.; Hu, Q.; Zhu, P.; Ma, P. Rule learning for classification based on neighborhood covering reduction. *Inf. Sci.* **2011**, *181*, 5457–5467. [CrossRef]
8. Zhang, X.; Li, J.; Li, W. A new mechanism of rule acquisition based on covering rough sets. *Appl. Intell.* **2022**, 1–13. [CrossRef]
9. Johnson, J.A.; Liu, M.; Chen, H. Unification of knowledge discovery and data mining using rough sets approach in a real-world application. *Rough Sets Curr. Trends Comput.* **2001**, *2005*, 330–337.

10. Wu, H.; Liu, G. The relationships between topologies and generalized rough sets. *Int. J. Approx. Reason.* **2020**, *119*, 313–324. [CrossRef]
11. Alcantud, J.C.R. Revealed indifference and models of choice behavior. *J. Math. Psychol.* **2002**, *46*, 418–430. [CrossRef]
12. Chen, Y.; Miao, D.; Wang, R.; Wu, K. A rough set approach to feature selection based on power set tree. *Knowl.-Based Syst.* **2011**, *24*, 275–281. [CrossRef]
13. Javidi, M.; Esk, A.S. Streamwise feature selection: A rough set method. *Int. J. Mach. Learn. Cybern.* **2018**, *9*, 667–676. [CrossRef]
14. Cai, Z.; Zhu, W. Multi-label feature selection via feature manifold learning and sparsity regularization. *Int. J. Mach. Learn. Cybern.* **2018**, *9*, 1321–1334. [CrossRef]
15. Luce, R.D. Semiorders and a theory of utility discrimination. *Econometrica* **1956**, *24*, 178–191. [CrossRef]
16. Wang, C.; He, Q.; Shao, M.; Xu, Y.; Hu, Q. A unified information measure for general binary relations. *Knowl.-Based Syst.* **2017**, *135*, 18–28. [CrossRef]
17. Zhu, W. Generalized rough sets based on relations. *Inf. Sci.* **2007**, *177*, 4997–5001. [CrossRef]
18. Zhu, W.; Wang, F. Reduction and axiomatization of covering generalized rough sets. *Inf. Sci.* **2003**, *152*, 217–230. [CrossRef]
19. Zhu, W. Relationship among basic concepts in covering-based rough sets. *Inf. Sci.* **2009**, *179*, 2478–2486. [CrossRef]
20. Yang, B.; Hu, B. On some types of fuzzy covering-based rough sets. *Fuzzy Sets Syst.* **2017** *312*, 36–65. [CrossRef]
21. Wang, C.; He, Q.; Wang, X.; Chen, D.; Qian, Y. Feature selection based on neighborhood discrimination index. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 2986–2999. [CrossRef]
22. Gao, N.; Li, Q.; Han, H.; Li, Z. Axiomatic approaches to rough approximation operators via ideal on a complete completely distributive lattice. *Soft Comput.* **2018**, *22*, 2329–2339. [CrossRef]
23. Wang, J.; Zhang, X.; Liu, C. Grained matrix and complementary matrix: Novel methods for computing information descriptions in covering approximation spaces. *Inf. Sci.* **2022**, *591*, 68–87. [CrossRef]
24. Liang, D.; Xu, Z.; Liu, D. Three-way decisions with intuitionistic fuzzy decision-theoretic rough sets based on point operators. *Inf. Sci.* **2017**, *375*, 183–201. [CrossRef]
25. Zhao, Z. On some types of covering rough sets from topological points of view. *Int. J. Approx. Reason.* **2016**, *68*, 1–14. [CrossRef]
26. Chiaselotti, G.; Ciucci, D.; Gentile, T. Rough set theory and digraphs. *Fundam. Inform.* **2017**, *153*, 291–325. [CrossRef]
27. Ali, Z.; Mahmood, T.; Ullah, K.; Khan, Q. Einstein geometric aggregation operators using a novel complex interval-valued pythagorean fuzzy setting with application in green supplier chain management. *Rep. Mech. Eng.* **2021**, *2*, 105–134. [CrossRef]
28. Sharma, H.K.; Singh, A.; Yadav, D.; Kar, S. Criteria selection and decision making of hotels using dominance based rough set theory. *Oper. Res. Eng. Sci. Theory Appl.* **2022**, *5*, 41–55. [CrossRef]
29. Chattopadhyay, R.; Das, P.P.; Chakraborty, S. Development of a rough-MABAC-DoE-based metamodel for supplier selection in an iron and steel industry. *Oper. Res. Eng. Sci. Theory Appl.* **2022**, *5*, 20–40. [CrossRef]
30. Liu, Y.; Zheng, L.; Xiu, Y.; Yin, H.; Zhao, S.; Wang, X.; Chen, H.; Li, C. Discernibility matrix based incremental feature selection on fused decision tables. *Int. J. Approx. Reason.* **2020**, *118*, 1–26. [CrossRef]
31. Yao, E.; Li, D.; Zhai, Y.; Zhang, C. Multi-label feature selection based on relative discernibility pair matrix. *IEEE Trans. Fuzzy Syst.* **2021**. [CrossRef]
32. Fan, X.; Chen, Q.; Qiao, Z.; Wang, C.; Ten, M. Attribute reduction for multi-label classification based on labels of positive region. *Soft Comput.* **2020**, *24*, 14039–14049. [CrossRef]
33. Ma, Y.; Luo, X.; Li, X.; Bao, Z. Selection of Rich Model Steganalysis Features Based on Decision Rough Set α-Positive Region Reduction. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *29*, 336–350. [CrossRef]
34. Zhang, X.; Mei, C.; Chen, D.; Zhang, Y.; Li, J. Active incremental feature selection using a fuzzy-rough-set-based information entropy. *IEEE Trans. Fuzzy Syst.* **2020**, *28*, 901–915. [CrossRef]
35. Xu, J.; Yuan, M.; Ma, Y. Feature selection using self-information and entropy-based uncertainty measure for fuzzy neighborhood rough set. *Complex Intell. Syst.* **2022**, *8*, 287–305. [CrossRef]
36. Zhang, X.; Wang, J.; Zhan, J.; Dai, J. Fuzzy measures and choquet integrals based on fuzzy covering rough sets. *IEEE Trans. Fuzzy Syst.* **2021**. [CrossRef]
37. Wang, J.; Zhang, X.; Yao, Y. Matrix approach for fuzzy description reduction and group decision-making with fuzzy β-covering. *Inf. Sci.* **2022**, *597*, 53–85. [CrossRef]
38. Komorowski, J.; Pawlak, Z.; Polwski, L.; Skowron, A. Rough sets: A tutorial. In *Rough Fuzzy Hybridization, A New Trend in Decision Making*; Pal, S., Skowron, K., Eds.; Springer: Berlin/Heidelberg, Germany, 1999.
39. Clark, P.E.; Niblett, T. The CN2 induction algorithm. *Mach. Learn.* **1986**, *3*, 261–284. [CrossRef]
40. Grzymala-Busse, J.W. A new version of the rule induction system LERS. *Fundam. Inform.* **1997**, *31*, 27–39. [CrossRef]
41. Yao, Y. Constructive and algebraic methods of theory of rough sets. *Inf. Sci.* **1998**, *109*, 21–47. [CrossRef]
42. Grabisch, M. *K*-order additive discrete FMs and their representation. *Fuzzy Sets Syst.* **1997**, *92*, 167–189. [CrossRef]
43. Sugeno, M. Theory of Fuzzy Integrals and Its Applications. Ph.D. Thesis, Tokyo Institute of Technology, Tokyo, Japan, 1974.
44. Grabisch, M. Fuzzy integral in multicriteria decision making. *Fuzzy Sets Syst.* **1995**, *69*, 279–289. [CrossRef]
45. Choquet, C. Theory of capacities. *Ann. L'institut Fourier* **1953**, *5*, 131–295. [CrossRef]
46. Michalski, R.S.; Kaufman, K.; Wnek, J. *The AQ Family of Learning Programs: A Review of Recent Developments and an Exemplary Application*; Reports of Machine Learning and Inference Laboratory, George Mason University: Fairfax, VA, USA, 1991.

47. Wang, J.; Zhang, X. Matrix approaches for some issues about minimal and maximal descriptions in covering-based rough sets. *Int. J. Approx. Reason.* **2019**, *104*, 126–143. [CrossRef]
48. Zhan, J.; Jiang, H.; Yao, Y. Three-way multi-attribute decision-making based on outranking relations. *IEEE Trans. Fuzzy Syst.* **2021**, *29*, 2844–2858. [CrossRef]
49. Zhang, X.; Wang, J. Fuzzy β-covering approximation spaces. *Int. J. Approx. Reason.* **2020**, *126*, 27–47. [CrossRef]
50. Pop, H.F.; Sarbu, C. A powerful supervised fuzzy method: Characterization, authentication and traceability of roman pottery. *Stud. Univ.-Babes-Bolyai Chem.* **2022**, *67*, 61–74. [CrossRef]
51. Liang, R.; Zhang, X. Interval-valued pseudo overlap functions and application. *Axioms* **2022**, *11*, 216. [CrossRef]
52. Wen, X.; Zhang, X. Overlap functions based (multi-granulation) fuzzy rough sets and their applications in MCDM. *Symmetry* **2021**, *13*, 1779. [CrossRef]