



Computational Intelligence and Information Management

A relational perspective of attribute reduction in rough set-based data analysis

Tuan-Fang Fan^a, Churn-Jung Liao^{b,*}, Duen-Ren Liu^c^a Department of Computer Science and Information Engineering, National Penghu University, Penghu 880, Taiwan^b Institute of Information Science, Academia Sinica, Taipei 115, Taiwan^c Institute of Information Management, National Chiao-Tung University, Hsinchu 300, Taiwan

ARTICLE INFO

Article history:

Received 20 November 2009

Accepted 14 August 2010

Available online 20 August 2010

Keywords:

Rough sets

Decision analysis

Fuzzy sets

Attribute reduction

Relational information system

ABSTRACT

Attribute reduction is very important in rough set-based data analysis (RSDA) because it can be used to simplify the induced decision rules without reducing the classification accuracy. The notion of reduct plays a key role in rough set-based attribute reduction. In rough set theory, a reduct is generally defined as a minimal subset of attributes that can classify the same domain of objects as unambiguously as the original set of attributes. Nevertheless, from a relational perspective, RSDA relies on a kind of dependency principle. That is, the relationship between the class labels of a pair of objects depends on component-wise comparison of their condition attributes. The larger the number of condition attributes compared, the greater the probability that the dependency will hold. Thus, elimination of condition attributes may cause more object pairs to violate the dependency principle. Based on this observation, a reduct can be defined alternatively as a minimal subset of attributes that does not increase the number of objects violating the dependency principle. While the alternative definition coincides with the original one in ordinary RSDA, it is more easily generalized to cases of fuzzy RSDA and relational data analysis.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The rough set theory proposed by Pawlak (1982) provides an effective tool for extracting knowledge from data tables. As noted in Pawlak (1991), knowledge is deep-seated in the classification capabilities of human beings. A classification is simply a partition of a universe. Thus, in rough set theory, objects are partitioned into equivalence classes based on their attribute-values, which are essentially functional information associated with the objects. Many databases only contain functional information about objects; however, data about the relationships between objects has become increasingly important in decision analysis. A remarkable example is social network analysis, in which the principal types of data are attribute data and relational data.

To represent attribute data, a data table in rough set theory consists of a set of objects and a set of attributes, where each attribute is viewed as a function from the set of objects to the domain of values of the attribute. Hence, such data tables are also called *functional information systems* (FIS), and rough set theory can be viewed as a theory of *functional granulation*. Recently, granulation based on relational information between objects, called *relational granulation*, was proposed by Liao and Lin (2005). To facilitate

the study of relational granulation, it is necessary to represent and reason about data in *relational information systems* (RIS) (Fan et al., 2006).

Interestingly, several RISs can usually be derived by comparing the attribute-values of objects in an FIS, thereby providing a relational perspective on the original FIS. For example, when rough set theory is applied to multi-criteria decision analysis (MCDA), it is crucial that preference-ordered attribute domains and decision classes be considered (Greco et al., 1999, 2001, 2008, 2010, Slowinski et al., 2009, Fortemps et al., 2008). The original rough set theory cannot handle inconsistencies arising from violations of the dominance principle due to its use of the indiscernibility relation. Therefore, in the dominance-based rough set approach (DRSA), the indiscernibility relation is replaced by a dominance relation to solve the multi-criteria sorting problem; and the data table is replaced by a pairwise comparison table to solve multi-criteria choice and ranking problems. A pairwise comparison table is actually a kind of RIS derived from the original multi-criteria decision data, which is usually represented as an FIS.

Rough set-based data analysis (RSDA) addresses the fundamental issue of attribute reduction in information systems. By using attribute reduction techniques, the decision rules extracted from an information system can be simplified without reducing the accuracy of the classification. The key notions in rough set-based attribute reduction are *reduct* and *core*. Intuitively, a reduct is an essential part of an information system that suffices to define all

* Corresponding author. Tel.: +886 2 27883799.

E-mail addresses: dfan@npu.edu.tw (T.-F. Fan), liaucj@iis.sinica.edu.tw (C.-J. Liao), dliu@iim.nctu.edu.tw (D.-R. Liu).

the basic categories occurring in the system; whereas the core is, in a certain sense, the most important part of the system. Traditionally, reduct is defined with respect to FISs in rough set theory. However, from the relational perspective of FIS mentioned above, RSDA actually relies on a kind of dependency principle. That is, the relationship between the class labels of a pair of objects depends on component-wise comparison of their condition attributes. The larger the number of condition attributes compared, the greater the probability that the dependency will hold. Thus, eliminating condition attributes may cause more object pairs to violate the dependency principle. Based on this observation, a reduct can be defined alternatively as a minimal subset of attributes that does not increase the number of objects violating the dependency principle. While the alternative definition coincides with the original one in ordinary RSDA, it can be easily generalized to cases of fuzzy RSDA and relational data analysis. In this paper, we elaborate on the relational perspective of attribute reduction in different kinds of information systems.

The remainder of this paper is organized as follows. In Section 2, we review several variants of information systems, including FIS and RIS. In Section 3, we present a general framework of attribute reduction from the relational perspective of information systems. In Section 4, we instantiate the framework to different kinds of information systems. Section 5 contains some concluding remarks.

2. Information systems

Information systems are fundamental to rough set theory (Pawlak, 1991). In this section, we review several variants of FIS used in rough set theory and the RIS proposed by Fan et al. (2006).

2.1. Functional information systems

In data mining problems, data is usually provided in the form of a data table, which is formally defined as an attribute-value information system, and taken as the basis of the approximation space in rough set theory (Pawlak, 1991). To emphasize the fact that each attribute in an attribute-value system is associated with a function on the set of objects, we call such systems *functional information systems*.

Definition 1. A functional information system (FIS)¹ is a quadruple

$$T_f = (U, A, \{V_i | i \in A\}, \{f_i | i \in A\}),$$

where U is a nonempty set, called the universe; A is a nonempty finite set of attributes; for each $i \in A$, V_i is the domain of values for i ; and for each $i \in A$, $f_i : U \rightarrow V_i$ is a total function.

In an FIS, the information about an object consists of the values of its attributes. Thus, given a subset of attributes $B \subseteq A$, we can define the information function associated with B as $Inf_B : U \rightarrow \prod_{i \in B} V_i$,

$$Inf_B(x) = (f_i(x))_{i \in B}. \tag{1}$$

Example 1. Let us consider the problem of ranking scientific journals. Assume that Table 1 is an FIS containing data about six journals, whose condition attributes are impact factor and citation half-life and decision attribute is the rank. Thus, in this FIS, we have $U = \{1, 2, \dots, 6\}$, $A = \{1, 2, d\}$, $V_i = [0, 10]$ for $i = 1, 2$, and $V_d = \{a, b, c\}$, and f_i is specified in Table 1.

Table 1
An FIS of the data about 6 journals.

U	A		
	f_1	f_2	f_d
1	3.5	5.5	a
2	3.5	5.5	a
3	2.0	1.5	a
4	2.0	1.5	b
5	2.0	1.0	a
6	2.0	1.0	b

2.2. Relational information systems

Though much information associated with individual objects is given in a functional form, it is sometimes more natural to represent such information in a relational form. For example, in a demographic database, it is more natural to represent the parent-child relationship as a relation between individuals, instead of an attribute of the parent or the child. In some cases, it may be necessary to use relational information simply because the exact values of some attributes are not available. For example, we may not know the exact ages of two individuals, but we do know which one is older. These considerations motivate the following definition of an alternative kind of information system called an RIS.

Definition 2. A relational information system (RIS) is a quadruple

$$T_r = (U, A, \{H_i | i \in A\}, \{r_i | i \in A\}),$$

where U and A are defined as above; for each $i \in A$, H_i is a set of relational indicators; and for each $i \in A$, $r_i : U \times U \rightarrow H_i$ is a total function.

A relational indicator in H_i is used to indicate the extent or degree to which two objects are related according to an attribute i . Thus, $r_i(x, y)$ denotes the extent to which x is related to y on the attribute i . If $H_i = \{0, 1\}$, then, for any $x, y \in U$, x is said to be i -related to y iff $r_i(x, y) = 1$. In most cases, several RISs can be derived from a given FIS by comparing the attribute-values of objects in the FIS.

Example 2. Continuing with Example 1, assume that we are interested in the pairwise comparison of the journals, instead of their real attribute values. Then, we may derive an RIS $T_r = (U, A, \{H_i | i \in A\}, \{r_i | i \in A\})$ from the original FIS, where U and A are defined as in Example 1, $H_i = \{0, 1\}$, and $r_i : U \times U \rightarrow \{0, 1\}$ is defined by

$$r_i(x, y) = 1 \iff f_i(x) \geq f_i(y),$$

for $i = 1, 2$ and $r_d(x, y) = 1$ iff $x = a$ or $y = c$ or $x = y = b$.

Mathematically, an RIS can be viewed as a special case of an FIS if we consider the Cartesian space $U \times U$ as a universe in itself. However, such a viewpoint ignores a subtle difference between FIS and RIS. That is, the objects in the universe of an FIS are supposed to be independent entities, whereas the pairs in the Cartesian space of an RIS may be inter-dependent. For example, if r_i is a symmetric relation, then the pairs (x, y) and (y, x) should have the same indicator value with respect to r_i in the RIS, while x and y are totally distinct objects.

On the other hand, an FIS $T_f = (U, A, \{V_i | i \in A\}, \{f_i | i \in A\})$ can be seen as a special case of an RIS $T_r = (U, A, \{H_i | i \in A\}, \{r_i | i \in A\})$ if we take $H_i = V_i \cup \{-\}$ and define $r_i(x, x) = f_i(x)$ and $r_i(x, y) = -$ if $x \neq y$. While this transformation seems somewhat trivial, an interesting fact is that more natural RIS may arise from a given FIS if the domains of attributes possess some kind of structure. As shown in Example 2, the RIS arises naturally from the original FIS due to the order structure of the domains of attributes. Indeed, most

¹ Originally called information systems, data tables, knowledge representation systems, or attribute-value systems in rough set theory.

variants of the information systems considered in this paper exhibit such a structure, and we actually utilize the derived RIS to define the general framework of attribute reduction. However, we note that the derived RIS is usually secondary in the sense that we can not restore the original FIS from the derived RIS.² For example, even though we know the relative order of the impact factors of the six journals in Example 2, we can not obtain their real values from the RIS.

2.3. Variants of information systems

In the definition of primitive FIS, the domain V_i of each attribute is simply a set of values without any structure. However, in many practical applications, natural structures are usually imposed on the domains of attributes. This results in many variants of the primitive FIS. We review the main variants in this subsection.

2.3.1. Preference-ordered decision table

For MCDA problems, each object in a decision table can be seen as a sample decision, and each condition attribute can be regarded as a criterion for the decision. Since a criterion's domain of values is usually ordered according to the decision-maker's preferences, each V_i is endowed with a binary relation \succeq_i . Thus, a *preference-ordered decision table* (PODT) is an FIS $T_f = (U, A, \{(V_i, \succeq_i) | i \in A\}, \{f_i | i \in A\})$. The relation \succeq_i is called a *weak preference relation* or *outranking* on V_i , and represents a preference over the domain of values of the criterion i (Slowinski et al., 2002). For $x, y \in U$, $f_i(x) \succeq_i f_i(y)$ means that “ x is at least as good as y with respect to criterion i ”. The weak preference relation \succeq_i is supposed to be a complete preorder, i.e., a complete, reflexive, and transitive relation. In addition, we assume that the domain of the decision attribute is a finite set $V_d = \{1, 2, \dots, n\}$ such that r is strictly preferred to s if $r > s$ for any $r, s \in V_d$. Thus, the weak preference relation \succeq_d is defined as $r \succeq_d s$ iff $r \geq s$; consequently, \succeq_d is a total order. Intuitively, for a condition criterion i and an object x , $f_i(x)$ denotes the *evaluation* of the object with respect to the criterion i ; and for the decision attribute d , $f_d(x)$ represents the *assignment* of x to a decision label in V_d .

2.3.2. Uncertain and multi-valued information systems

When each attribute's domain of values is the powerset of another base domain, we can model both uncertain and multi-valued information systems. Let us consider an FIS $T_f = (U, A, \{2^{V_i} | i \in A\}, \{f_i | i \in A\})$. It is an *uncertain information system* (UIS) if it is interpreted disjunctively (Kryszkiewicz and Rybiński, 1996a,b, 1998). In other words, we do not know the exact value of attribute i of the object x , but we do know that the value is in the set $f_i(x)$. By contrast, T_f is a *multi-valued information system* (MVIS) if it is interpreted conjunctively. This means that *all* the values in $f_i(x)$ are deemed to be the values of attribute i of object x . For example, if i is the course(s) taken by x , then in an MVIS, $f_i(x) = \{\text{Algebra, Algorithm, PL}\}$ means that x takes all these courses.

2.3.3. Proximity-based and metric-based information systems

When each V_i is endowed with a (crisp) reflexive and symmetric relation \sim_i , we call $T_f = (U, A, \{(V_i, \sim_i) | i \in A\}, \{f_i | i \in A\})$ a *proximity-based information system* (PIS). For example, if the attribute i denotes a location, there may be a (qualitative) nearness relation between different locations in V_i . On the other hand, an FIS $T_f = (U, A, \{(V_i, \delta_i) | i \in A\}, \{f_i | i \in A\})$ is called a *metric-based information system* (MIS) if there is a metric $\delta_i : V_i \times V_i \rightarrow [0, 1]$ between ele-

Table 2
Example of an UIS and a POUDT.

U	A		
	f_1	f_2	f_d
1	[46, 50]	[48, 52]	4
2	[44, 48]	[48, 50]	4
3	[45, 52]	44	3
4	26	[28, 35]	3
5	30	[26, 32]	2
6	[24, 27]	33	[2, 3]
7	16	[10, 16]	[2, 3]
8	24	10	[2, 3]

ments of V_i . A typical example is the case where V_i is the Euclidean space and δ_i is the distance metric.

2.3.4. Preference-ordered uncertain decision table

Although the PODT can represent multi-criteria decision cases effectively, it inherits the restriction of the classical FIS, which means that uncertain information can not be represented. As UISs generalize FISs, PODTs can be generalized to *preference-ordered uncertain decision tables* (POUDTs). Formally, a POUDT is an FIS $T_f = (U, A, \{2^{V_i} | i \in A\}, \{\succeq_i | i \in A\}, \{f_i | i \in A\})$, where, as in the case of a UIS, each domain of attributes is the power set of some base domain; and as in the case of a PODT, each base domain V_i is endowed with a weak preference relation \succeq_i . The intuition about a POUDT is that the evaluation of a criterion i for an object x belongs to $f_i(x)$, although the evaluation is not known exactly. Furthermore, we assume that for each criterion i , the space $V_i \times V_i$ is endowed with a uniform measure μ_i . Thus, for each bounded subset $F \subseteq V_i \times V_i$, $\mu_i(F)$ is a non-negative real number. When V_i is a finite set, we take $\mu_i(F)$ as the cardinality of F ; and when V_i is a real interval, we take $\mu_i(F)$ as the area of F . The measure is used to compare two subsets of V_i based on the weak preference relation \succeq_i .

Example 3. Let us consider an example extracted from (Dembczynski et al. (2009)). Table 2 is a UIS containing eight objects with interval evaluations and assignments. Since an interval is a subset of the domain, it is in fact an instance of a UIS. Furthermore, a natural ordering is endowed with the underlying domain of criteria, i.e., the set of real numbers. In other words, we assume that $r \succeq_i s$ iff $r \geq s$. Therefore, the table is also an instance of a POUDT. This is simply a generic example for illustrative purposes. However, it is straightforward to convert it into a real-world example by assigning appropriate interpretations to the criteria. For instance, in the financial domain, we can consider the data in Table 2 as evaluations of mortgage applications. Then, the criteria may represent the income of the applicants and the value of the target houses. Moreover, due to privacy concerns, it has been suggested that precise values should be replaced by uncertain ones for some sensitive attributes like personal income when a data table is released for analysis (Sweeney, 2002a,b, Wang et al., 2004, 2006, 2007). Thus, this kind of POUDT may arise naturally in such context.

3. Dependency constraints and attribute reduction in information systems

3.1. Dependency constraints and the consistency of information systems

Let us consider an FIS $T_f = (U, A, \{V_i | i \in A\}, \{f_i | i \in A\})$, where A can be partitioned into a set of condition attributes C and a decision attribute d , i.e., $A = C \cup \{d\}$. Such an FIS is also called a decision

² In some cases, we can obtain an FIS isomorphic to the original one from the derived RIS. See Fan et al. (2006) for further details.

table. Hereafter, we only consider decision tables; thus, when we discuss an FIS, we assume that its set of attributes is $A = C \cup \{d\}$. We further assume that there is a common set of relational indicators H such that, for each $i \in A$, there exists a comparison operator $\theta_i : f_i(U) \times f_i(U) \rightarrow H$. Note that we intentionally define the domain of θ_i as $f_i(U) \times f_i(U)$ instead of $V_i \times V_i$ because, in some information systems studied in this paper, $f_i(U)$ is a subset of 2^{V_i} instead of V_i . Let $\Theta = (\theta_i)_{i \in A}$; then, Θ can induce a transformation from the given FIS to an RIS $\Theta(T_f) = (U, A, H, \{r_i | i \in A\})$, where $r_i(x, y) = f_i(x)\theta_i f_i(y)$. By convention, we use the infix notation instead of the prefix notation for θ_i . As in the case of FIS, we define the information function associated with a subset of attributes B for the system $\Theta(T_f)$ as $\text{Inf}_B : U^2 \rightarrow H^{|B|}$,

$$\text{Inf}_B(x, y) = (f_i(x)\theta_i f_i(y))_{i \in B}. \tag{2}$$

To aggregate the results of comparing different attributes, we need an associative and commutative aggregation operator $\otimes : H^2 \rightarrow H$ on the set of relational indicators. Since \otimes is associative and commutative, we can write $((h_1 \otimes h_2) \cdot \dots \otimes h_k)$ as $\otimes(h_1, \dots, h_k)$. Let $\Rightarrow \subseteq H \times H$ be a binary implication relation on H , and let $B \subseteq C$ be a subset of condition attributes. Then, we can define a binary relation $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d) \subseteq U \times U$ as follows:

$$\{(x, y) | \otimes \text{Inf}_B(x, y) \Rightarrow r_d(x, y), x \in U, y \in U\}, \tag{3}$$

where we use the infix notation for the implication relation \Rightarrow . We say that a pair of objects (x, y) satisfies the $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)$ -dependency constraint if $(x, y) \in \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)$. For a binary relation $R \subseteq U \times U$, let R^+ , R^- , and R^{cs} denote $\{x | \forall y(x, y) \in R\}$, $\{x | \forall y(y, x) \in R\}$, and $R^+ \cap R^-$ respectively. Then, an object x is $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)$ -consistent if $x \in \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)^{\text{cs}}$.

Example 4. Continuing with Examples 1 and 2, let us assume that Θ is the transformation that changes the FIS in Example 1 to the RIS in Example 2. Then, $H = \{0, 1\}$ and $B = \{1, 2\}$ imply that $\text{Inf}_B(x, y) = \{(x, y) | x, y \in U, x \leq y\}$. Consequently, $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d) = U \times U - \{(4, 5)\}$, $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)^+ = \{1, 2, 3, 5, 6\}$, $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)^- = \{1, 2, 3, 4, 6\}$, and the set of consistent objects is $\{1, 2, 3, 6\}$.

3.2. Attribute reduction

Let B be a subset of attributes and let i be an attribute in B . Then, we say that i is *relationally dispensable* in B (with respect to the dependency constraint $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)$)³ if $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B - \{i\}, d) = \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)$; otherwise, i is *relationally indispensable* in B . The set of attributes B is *relationally independent* if each $i \in B$ is relationally indispensable in B ; otherwise B is *relationally dependent*. The set of attributes $B \subseteq C$ is a *relation-based reduct* of the FIS T_f if B is relationally independent and $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d) = \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(C, d)$. The set of all relationally indispensable attributes in C is called the *relation-based core* of T_f , and is denoted by $\text{RCORE}(T_f)$. Analogously, we can define an *object-based reduct and core* by replacing $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B - \{i\}, d) = \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)$ with $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B - \{i\}, d)^{\text{cs}} = \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)^{\text{cs}}$ and $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d) = \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(C, d)$ with $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}(B, d)^{\text{cs}} = \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(C, d)^{\text{cs}}$ in the above definitions. The object-based core of T_f is denoted by $\text{CCORE}(T_f)$. An obvious relation exists between these two types of reducts and cores.

Proposition 1

1. For any relation-based reduct B , there must exist an object-based reduct B' such that $B' \subseteq B$. In other words, every relation-based

reduct can be shrunk to an object-based reduct. By contrast, not every object-based reduct is a subset of some relation-based reduct.
 2. $\text{CCORE}(T_f) \subseteq \text{RCORE}(T_f)$

Before proceeding, we must consider two assumptions regarding the aggregation operator \otimes and the implication relation \Rightarrow , called *monotonicity* and *decomposability* respectively.

- (Mon) For all h_1, h_2 , and $h_3 \in H$, if $h_1 \Rightarrow h_3$, then $h_1 \otimes h_2 \Rightarrow h_3$.
- (Decom) For all h_1, h_2 , and $h_3 \in H$, if $h_1 \otimes h_2 \Rightarrow h_3$, then $h_1 \Rightarrow h_3$ or $h_2 \Rightarrow h_3$.

The monotonicity assumption is mandatory for the definition of a reduct; indeed, all real cases that we consider in the next section satisfy this assumption. On the other hand, decomposability does not hold for all the cases considered in the next section. If the assumption does hold, we can use the discernibility matrix proposed by Skowron and Rauszer (1991) to find the reduct and core of a system. The *relation-based discernibility matrix* of T_f is defined as $\mathcal{D}_{T_f} : U \times U \rightarrow 2^C$:

$$\mathcal{D}_{T_f}(x, y) = \begin{cases} \emptyset, & \text{if } \forall h \in H, h \Rightarrow r_d(x, y) \\ \{i \in C | r_i(x, y) \Rightarrow r_d(x, y)\}, & \text{otherwise} \end{cases}$$

Then, we have the following proposition.

Proposition 2. If \otimes and \Rightarrow satisfy both the monotonicity and decomposability assumptions, then

1. an attribute $i \in \text{RCORE}(T_f)$ iff there exist $x, y \in U$ such that $\mathcal{D}_{T_f}(x, y) = \{i\}$ and
2. a subset of attributes B is a relation-based reduct of T_f iff $\bigwedge B$ is a prime implicant of the propositional formula $\bigwedge \{\bigvee \mathcal{D}_{T_f}(x, y) | x, y \in U, \mathcal{D}_{T_f}(x, y) \neq \emptyset\}$.

An analogous result holds for an object-based reduct and core if we replace the definition of \mathcal{D}_{T_f} with that of the *object-based discernibility matrix* as follows:

$$\mathcal{D}_{T_f}(x, y) = \begin{cases} \emptyset, & \text{if } \forall h \in H, h \Rightarrow r_d(x, y) \\ & \text{or } x, y \notin \mathfrak{D}_{\Theta, \otimes, \Rightarrow}(C, d)^{\text{cs}} \\ \{i \in C | r_i(x, y) \Rightarrow r_d(x, y)\}, & \text{otherwise} \end{cases}$$

4. Case studies

In the general framework presented in Section 3, there are three parameters for the dependency constraint, i.e., Θ (including the range H), \otimes , and \Rightarrow . We now show that different RSDA solutions for attribute reduction can be derived through the instantiations of the three parameters. The cases considered in this section can be divided into two types of systems based on the instantiations of H , \otimes , and \Rightarrow :

1. The crisp type: $H = \{0, 1\}$, \otimes is the Boolean conjunction \wedge , and \Rightarrow is the material implication of Boolean logic \rightarrow . For this type of system, we abbreviate the dependency constraint $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}$ as \mathfrak{C}_{Θ} . Note that the decomposability assumption always holds for this type of system, so the discernibility matrix method can be applied.
2. The fuzzy type: $H = [0, 1]$, \otimes is a t-norm,⁴ and \Rightarrow is the \leq relation on the unit interval. For this type of system, we abbreviate the dependency constraint $\mathfrak{D}_{\Theta, \otimes, \Rightarrow}$ as \mathfrak{F}_{Θ} . The decomposability assumption holds for this type of system when the t-norm

³ Hereafter, we omit the qualifier “with respect to the dependency constraint ...”, when it is clear from the context.

⁴ For the properties of t-norms, readers may refer to any standard reference on fuzzy logic, such as (Hájek, 1998).

$\otimes = \min$. Thus, in the case of $\otimes = \min$, the discernibility matrix method can be applied.

4.1. Classical rough set approach

The basic construct of rough set theory is an approximation space, which is defined as a pair (U, R) , where U is a finite universe and $R \subseteq U \times U$ is an equivalence relation on U . We write an equivalence class of R as $[x]_R$ if it contains the element x . For an approximation space (U, R) and an arbitrary concept $X \subseteq U$, we are interested in defining X based on the equivalence classes of R . We say that X is R -definable if X is a union of some R -equivalence classes; otherwise X is R -undefinable. The R -definable concepts are also called R -exact sets, whereas R -undefinable concepts are said to be R -inexact or R -rough sets. A rough set can be approximated from below and above by two exact sets. The lower and upper approximations of X are denoted by $\underline{R}X$ and $\overline{R}X$ respectively, and defined as follows:

$$\underline{R}X = \{x \in U \mid [x]_R \subseteq X\},$$

$$\overline{R}X = \{x \in U \mid [x]_R \cap X \neq \emptyset\}.$$

The main concept of the classical rough set approach (CRSA) is based on the definition of the indiscernibility relation. Let $T_f = (U, A, \{V_i \mid i \in A\}, \{f_i \mid i \in A\})$ be a decision table and $B \subseteq A$ be a subset of attributes. Then, the indiscernibility relation with respect to B is defined on U as follows:

$$\text{ind}(B) = \{(x, y) \mid x, y \in U, f_i(x) = f_i(y) \forall i \in B\}.$$

In other words, x and y are B -indiscernible if they have the same values with respect to all the attributes in B . Consequently, for each $B \subseteq A$, $(U, \text{ind}(B))$ is an approximation space in rough set theory. Thus, we write $\underline{B}X$ and $\overline{B}X$ for $\text{ind}(B)\underline{X}$ and $\text{ind}(B)\overline{X}$ respectively. Let U/d denote the family of equivalence classes of $\text{ind}(\{d\})$; then, the B -positive region of T_f , denoted by $\text{POS}_B(T_f)$, is defined as

$$\text{POS}_B(T_f) = \bigcup_{X \in U/d} \underline{B}X.$$

In CRSA, an attribute $i \in B$ is dispensable if $\text{POS}_{B-\{i\}}(T_f) = \text{POS}_B(T_f)$, and the definition of independence is as above. Then, a set of attributes $B \subseteq C$ is a reduct of T_f if B is independent and $\text{POS}_B(T_f) = \text{POS}_C(T_f)$. The core of T_f is simply the set of all indispensable attributes in C .

Let B be a subset of C . Then, an object x is B -consistent (with respect to the decision attribute d) if $[x]_{\text{ind}(B)} \subseteq [x]_{\text{ind}(\{d\})}$; otherwise, x is B -inconsistent. In other words, x is a B -consistent object in a decision table if, for all $y \in U$, (x, y) satisfies the following B -indiscernibility principle:

$$(x, y) \in \text{ind}(B) \Rightarrow (x, y) \in \text{ind}(\{d\}).$$

In other words, the objects that have the same values as x in terms of the condition attributes should have the same decision class assignment as x .

Let $\Theta_{=}$ denote $(=)_{i \in A}$, where $=_i$ is the identity relation on V_i . Then, it is straightforward to prove the following lemma.

Lemma 1. For any subset of attributes B and any $x, y \in U$,

1. (x, y) satisfies the B -indiscernibility principle iff $(x, y) \in \Theta_{\otimes}(B, d)$ and
2. x is a B -consistent object iff $x \in \Theta_{\otimes}(B, d)^{\text{tr}}$ iff $x \in \text{POS}_B(T_f)$.

From this lemma, it is straightforward to derive the following theorem.

Theorem 1. A subset of attributes B is a reduct (resp. the core) of T_f in CRSA iff it is an object-based reduct (resp. the object-based core) of T_f with respect to the $\Theta_{\otimes}(C, d)$ dependency constraint.

Example 5. Continuing with Example 1, let C be the set of all condition attributes. Then, each entry (x, y) of the following matrix denotes $(f_1(x) = f_1(y), f_2(x) = f_2(y), f_d(x) = f_d(y))$.

	1	2	3	4	5	6
1	(1, 1, 1)	(1, 1, 1)	(0, 0, 1)	(0, 0, 0)	(0, 0, 1)	(0, 0, 0)
2	(1, 1, 1)	(1, 1, 1)	(0, 0, 1)	(0, 0, 0)	(0, 0, 1)	(0, 0, 0)
3	(0, 0, 1)	(0, 0, 1)	(1, 1, 1)	(1, 1, 0)	(1, 0, 1)	(1, 0, 0)
4	(0, 0, 0)	(0, 0, 0)	(1, 1, 0)	(1, 1, 1)	(1, 0, 0)	(1, 0, 1)
5	(0, 0, 1)	(0, 0, 1)	(1, 0, 1)	(1, 0, 0)	(1, 1, 1)	(1, 1, 0)
6	(0, 0, 0)	(0, 0, 0)	(1, 0, 0)	(1, 0, 1)	(1, 1, 0)	(1, 1, 1)

Thus, since \otimes is the Boolean conjunction and \Rightarrow is the material implication, we have $(x, y) \in \Theta_{\otimes}(C, d) = U^2 - \{(3, 4), (4, 3), (5, 6), (6, 5)\}$ iff (x, y) satisfies the C -indiscernibility principle and $\Theta_{\otimes}(C, d)^{\text{tr}} = \{1, 2\} = \text{POS}_C(T_f)$. Let $B = \{1\}$; then we can see that $\Theta_{\otimes}(B, d) = U^2 - \{(3, 4), (4, 3), (5, 6), (6, 5), (5, 4), (4, 5), (3, 6), (6, 3)\}$ and $\Theta_{\otimes}(B, d)^{\text{tr}} = \{1, 2\}$. Thus, B is an object-based reduct, but not a relation-based reduct. In the same way, we can see that $B' = \{2\}$ is both an object-based reduct and a relation-based reduct. Consequently, $\emptyset = \text{CCORE}(T_f) \subseteq \text{RCORE}(T_f) = B'$. Therefore, the example confirms Proposition 1.

Example 6. Continuing with Example 5, since the Boolean conjunction and implication satisfy monotonicity and decomposability assumptions, we can use the discernibility matrix method to find the reducts and the core. First, the relation-based discernibility matrix is constructed as follows⁵:

	1	2	3	4	5	6
1	\emptyset	\emptyset	\emptyset	$\{f_1, f_2\}$	\emptyset	$\{f_1, f_2\}$
2	\emptyset	\emptyset	\emptyset	$\{f_1, f_2\}$	\emptyset	$\{f_1, f_2\}$
3	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	$\{f_2\}$
4	$\{f_1, f_2\}$	$\{f_1, f_2\}$	\emptyset	\emptyset	$\{f_2\}$	\emptyset
5	\emptyset	\emptyset	\emptyset	$\{f_2\}$	\emptyset	\emptyset
6	$\{f_1, f_2\}$	$\{f_1, f_2\}$	$\{f_2\}$	\emptyset	\emptyset	\emptyset

Thus, $\text{RCORE}(T_f) = \{f_2\}$ because $\{f_2\}$ is the only singleton in the discernibility matrix, and $\{f_2\}$ is a relation-based reduct since f_2 is the only prime implicant of the propositional formula $\bigwedge \{\bigvee \mathcal{D}_{T_f}(x, y) \mid x, y \in U, \mathcal{D}_{T_f}(x, y) \neq \emptyset\} = (f_1 \vee f_2) \wedge f_2 = f_2$. Furthermore, since $\Theta_{\otimes}(C, d)^{\text{tr}} = \{1, 2\}$, the object-based discernibility matrix is constructed as follow:

	1	2	3	4	5	6
1	\emptyset	\emptyset	\emptyset	$\{f_1, f_2\}$	\emptyset	$\{f_1, f_2\}$
2	\emptyset	\emptyset	\emptyset	$\{f_1, f_2\}$	\emptyset	$\{f_1, f_2\}$
3	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset
4	$\{f_1, f_2\}$	$\{f_1, f_2\}$	\emptyset	\emptyset	\emptyset	\emptyset
5	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset
6	$\{f_1, f_2\}$	$\{f_1, f_2\}$	\emptyset	\emptyset	\emptyset	\emptyset

⁵ Note that, to avoid confusion, we intentionally write the attributes as f_i instead of simply i .

Thus, $CCORE(T_f) = \emptyset$ since no singleton appears in the matrix; and both $\{f_1\}$ and $\{f_2\}$ are object-based reducts because f_1 and f_2 are both prime implicants of the propositional formula $\bigwedge\{\bigvee \mathcal{D}_{T_f}(x,y)|x,y \in U, \mathcal{D}_{T_f}(x,y) \neq \emptyset\} = f_1 \vee f_2$.

4.2. Dominance-based rough set approach

In this subsection, we consider the dominance-based rough set approach (DRSA) proposed by Greco et al. (2001). Let $T_f = (U, A, \{(V_i, \succeq_i) | i \in A\}, \{f_i | i \in A\})$ be a PODT and let P be a subset of criteria. Then, we can define the P -dominance relation $D_P \subseteq U \times U$ as follows:

$$(x, y) \in D_P \iff f_i(x) \succeq_i f_i(y) \forall i \in P.$$

When $(x, y) \in D_P$, we say that x P -dominates y , and that y is P -dominated by x . We usually use the infix notation $x D_P y$ to denote $(x, y) \in D_P$. Although each \succeq_i is a complete preorder, the dominance relation may be a preorder. If $P = \{i\}$ is a singleton, we write D_i instead of $D_{\{i\}}$. The basic principle underlying DRSA is called the dominance principle. Let P denote a subset of condition criteria. Then, the dominance principle with respect to P can be expressed for $x, y \in U$ as follows:

$$x D_P y \Rightarrow x D_d y.$$

The principle states that if x P -dominates y (i.e., x is at least as good as y with respect to all criteria in P), then x should be assigned to a decision class at least as good as the class assigned to y .

In CRSA, a consistent object must satisfy the indiscernibility principle; however, in DRSA, we require adherence to the dominance principle. Thus, an object x is P -consistent in T_f if for all $y \in U$, we have

$$(x D_P y \Rightarrow x D_d y) \wedge (y D_P x \Rightarrow y D_d x);$$

otherwise, x is P -inconsistent. Note that the dominance principle implies the indiscernibility principle because of the reflexivity of the dominance relation and the antisymmetry of \succeq_d .

Given the dominance relation D_P , the P -dominating set and P -dominated set of x are defined, respectively, as

$$D_P^+(x) = \{y \in U | y D_P x\}$$

and

$$D_P^-(x) = \{y \in U | x D_P y\}.$$

In addition, for each $t \in V_d$, we define the decision class Cl_t as $\{x \in U | f_d(x) = t\}$. Then, the upward and downward unions of classes can be defined as

$$Cl_t^{\succeq} = \bigcup_{s \succeq t} Cl_s$$

and

$$Cl_t^{\preceq} = \bigcup_{s \preceq t} Cl_s,$$

respectively. Based on the P -dominating sets and P -dominated sets, we can define the P -lower and P -upper approximations of Cl_t^{\succeq} and Cl_t^{\preceq} for each $t \in V_d$ as follows:

$$\underline{P}(Cl_t^{\succeq}) = \{x \in U | D_P^+(x) \subseteq Cl_t^{\succeq}\},$$

$$\overline{P}(Cl_t^{\succeq}) = \{x \in U | D_P^-(x) \cap Cl_t^{\succeq} \neq \emptyset\},$$

$$\underline{P}(Cl_t^{\preceq}) = \{x \in U | D_P^-(x) \subseteq Cl_t^{\preceq}\},$$

$$\overline{P}(Cl_t^{\preceq}) = \{x \in U | D_P^+(x) \cap Cl_t^{\preceq} \neq \emptyset\}.$$

The P -boundaries of Cl_t^{\succeq} and Cl_t^{\preceq} are then defined as

$$Bn_P(Cl_t^{\succeq}) = \overline{P}(Cl_t^{\succeq}) - \underline{P}(Cl_t^{\succeq})$$

and

$$Bn_P(Cl_t^{\preceq}) = \overline{P}(Cl_t^{\preceq}) - \underline{P}(Cl_t^{\preceq}),$$

respectively. Let $\mathbf{CI} = \{Cl_t | t \in V_d\}$ denote the partition of the universe U into decision classes. Then, the quality of the approximation of the partition \mathbf{CI} based on the set of criteria P can be defined as the ratio

$$\gamma_P(\mathbf{CI}) = \frac{|U - (\bigcup_{t \in V_d} Bn_P(Cl_t^{\succeq}) \cup \bigcup_{t \in V_d} Bn_P(Cl_t^{\preceq}))|}{|U|}.$$

Note that $\gamma_P(\mathbf{CI})$ is equal to the ratio of P -consistent objects in the universe U .

In DRSA, a criterion $i \in P$ is dispensable if $\gamma_{P-\{i\}}(\mathbf{CI}) = \gamma_P(\mathbf{CI})$, and the definition of independence is as above. Then, a set of criteria $P \subseteq C$ is a reduct of T_f if P is independent and $\gamma_P(\mathbf{CI}) = \gamma_C(\mathbf{CI})$. The core of T_f is again the set of all indispensable attributes in C .

Let Θ_{\succeq} denote $(\succeq_i)_{i \in A}$. Then, we have the following lemma.

Lemma 2. For any subset of criteria P and any $x, y \in U$,

1. (x, y) satisfies the dominance principle with respect to P iff $(x, y) \in \mathfrak{C}_{\Theta_{\succeq}}(P, d)$ and
2. x is a P -consistent object iff $x \in \mathfrak{C}_{\Theta_{\succeq}}(P, d)^{\text{c}}$.

From this lemma, we can easily derive the following theorem

Theorem 2. A subset of criteria P is a reduct (resp. the core) of T_f in DRSA iff it is an object-based reduct (resp. the object based core) of T_f with respect to the $\mathfrak{C}_{\Theta_{\succeq}}(C, d)$ dependency constraint.

4.3. Tolerance-based rough set approach

The tolerance-based rough set approach (TRSA) replaces the indiscernibility relation in CRSA with a tolerance relation, i.e., a reflexive and symmetric (but not necessarily transitive) binary relation (Skowron and Stepaniuk, 1996). TRSA can be applied to UIS or PIS. If T_f is a PIS $(U, A, \{(V_i, \sim_i) | i \in A\}, \{f_i | i \in A\})$, it is straightforward to obtain a similarity relation \sim_i for each $i \in A$. On the other hand, if T_f is a UIS $(U, A, \{2^{V_i} | i \in A\}, \{f_i | i \in A\})$, we can define $\sim_i \subseteq 2^{V_i} \times 2^{V_i}$ as $F \sim_i G$ iff $F \cap G \neq \emptyset$. In both cases, we have a relation \sim_i defined over the domain $f_i(U)$ for each $i \in A$.

Once the relation \sim_i is defined, a tolerance relation $sim(B) \subseteq U \times U$ can be derived for each subset of attributes B as follows:

$$sim(B) = \{(x, y) | x, y \in U, f_i(x) \sim_i f_i(y) \forall i \in B\}.$$

Consequently, for each $B \subseteq A$, $(U, sim(B))$ is a tolerance-based approximation space.

Let $[x]_{sim(B)} = \{y | (x, y) \in sim(B)\}$ denote the tolerance class of an object x with respect to the relation $sim(B)$. Then, the lower and upper approximations of any $X \subseteq U$ in the space $(U, sim(B))$, denoted by $\underline{B}^L X$ and $\overline{B}^L X$ respectively, are defined as in CRSA, except that the equivalence classes are replaced by tolerance classes. The B -positive region of T_f in TRSA, denoted by $POS_B^+(T_f)$, is defined as $POS_B^+(T_f) = \bigcup_{x \in U/d} \underline{B}^L X$. Hereafter, the notions of reduct and core are defined in the same way as in CRSA. Let us define the B -tolerance principle for $x, y \in U$ as follows:

$$(x, y) \in sim(B) \Rightarrow (x, y) \in ind(\{d\}).$$

Then, it is obvious that, in TRSA, the B -positive region is also the set of all B -consistent objects, where an object x is B -consistent if, for any object y , both (x, y) and (y, x) satisfy the B -tolerance principle.

Let us define Θ_{\sim} as $((\sim_i)_{i \in C} = d)$. Then, with an analogous argument to that used in the case of CRSA, we can derive the following lemma and theorem.

Lemma 3. For any subset of attributes B and any $x, y \in U$,

1. (x, y) satisfies the B -tolerance principle iff $(x, y) \in \mathfrak{C}_{\theta_{\sim}}(B, d)$;
2. x is a B -consistent object iff $x \in \mathfrak{C}_{\theta_{\sim}}(B, d)^{\text{cs}}$ iff $x \in \text{POS}_B^{\text{t}}(T_f)$.

Theorem 3. A subset of attributes B is a reduct (resp. the core) of T_f in TRSA iff it is an object-based reduct (resp. the object-based core) of T_f with respect to the $\mathfrak{C}_{\theta_{\sim}}(C, d)$ dependency constraint.

4.4. Fuzzy rough set approach

While TRSA deals with the crisp tolerance relation, the fuzzy similarity relation plays an important role in MIS and MVIS. Let $T_f = (U, A, \{(V_i, \delta_i) | i \in A\}, \{f_i | i \in A\})$ be an MIS. Then, we can define a fuzzy similarity relation $\approx_i : V_i \times V_i \rightarrow [0, 1]$ as $(v_1 \approx_i v_2) = 1 - \delta_i(v_1, v_2)$. On the other hand, if $T_f = (U, A, \{2^{V_i} | i \in A\}, \{f_i | i \in A\})$ is an MVIS such that each underlying domain V_i is finite, then we can define a fuzzy similarity relation $\approx_i : 2^{V_i} \times 2^{V_i} \rightarrow [0, 1]$ by using the Jaccard index (Jaccard, 1908)

$$F_{\approx_i}G = \frac{|F \cap G|}{|F \cup G|},$$

or the simple matching coefficient (Sokal and Sneath, 1963)

$$F_{\approx_i}G = \frac{|F \cap G| + |\bar{F} \cap \bar{G}|}{|V_i|}.$$

Once a fuzzy relation \approx_i over the domain $f_i(U)$ is given for each $i \in A$, a fuzzy similarity relation $f_{sim}(B) : U \times U \rightarrow [0, 1]$ can be derived for each subset of attributes B as follows:

$$f_{sim}(B)(x, y) = \otimes_{i \in B} f_i(x) \approx_i f_i(y).$$

Attribute reduction in the fuzzy rough set approach (FRSA) can be defined with the help of the fuzzy similarity relation. However, unlike in crisp cases, we do not define the reduct and core from the positive region of the system, since the region is a fuzzy subset of the universe. Instead, we adopt the following B -gradualness principle:

$$f_{sim}(B)(x, y) \leq f_{sim}(d)(x, y),$$

which means that the greater the similarity between the values of the condition attributes, the more similar the decisions will be. An object is B -fuzzily consistent if for any $y \in U$, both (x, y) and (y, x) satisfy the B -gradualness principle. The set of all B -fuzzily consistent objects of T_f (no matter whether it is an MVIS or an MIS) is denoted by $FCON_B(T_f)$. An attribute $i \in B$ is dispensable if $FCON_{B-\{i\}}(T_f) = FCON_B(T_f)$. Every minimal subset $B \subseteq C$ such that $FCON_B(T_f) = FCON_C(T_f)$ is a reduct of T_f , and the core of T_f is still the set of all indispensable attributes in C .

Let θ_{\approx} denote $(\approx_i)_{i \in A}$. Then, the following lemma and theorem can be derived:

Lemma 4. For any subset of attributes B and any $x, y \in U$,

1. (x, y) satisfies the B -gradualness principle iff $(x, y) \in \mathfrak{F}_{\theta_{\approx}}(B, d)$;
2. $x \in \mathfrak{F}_{\theta_{\approx}}(B, d)^{\text{cs}}$ iff $x \in FCON_B(T_f)$.

Theorem 4. A subset of attributes B is a reduct (resp. the core) of T_f in FRSA iff it is an object-based reduct (resp. the object-based core) of T_f with respect to the $\mathfrak{F}_{\theta_{\approx}}(C, d)$ dependency constraint.

4.5. Fuzzy dominance-based rough set approach

Because a POU DT is essentially a combination of a UIS and a PODT, it is appropriate to use the fuzzy dominance-based rough set approach (FDRSA) for decision analysis of POU DT. The objects

in a POU DT may have imprecise evaluations with respect to the condition criteria and imprecise assignments to the decision classes. Thus, the dominance relation between objects can not be determined with certainty. Instead, we define a degree of dominance between two objects with respect to each criterion i based on the associated measures μ_i .

Let $T_f = (U, A, \{2^{V_i} | i \in A\}, \{\succeq_i | i \in A\}, \{f_i | i \in A\})$ be a POU DT with the uniform measure μ_i being associated with $V_i \times V_i$ for $i \in A$. Then, we first define the fuzzy preference relation with respect to the criterion i as $\succsim_i : 2^{V_i} \times 2^{V_i} \rightarrow [0, 1]$ such that for all $F, G \subseteq V_i$,

$$F \succsim_i G = \frac{\mu_i(\{(v_1, v_2) | v_1 \succeq_i v_2, v_1 \in F, v_2 \in G\})}{\mu_i(F \times G)}.$$

Fig. 1 shows an example of computing the fuzzy preference relation, where the real evaluation of x with respect to criterion i , denoted by $s(x)$, is in a continuous interval $f_i(x) = [l_x, u_x]$. In this example, $f_i(x) \succsim_i f_i(y)$ is the ratio of the area of ABC over the area of $ABDE$, i.e., $\frac{u_x - l_y}{2(l_y - l_x)}$. Next, by applying the fuzzy preference relation on each criterion, we can define a fuzzy P -dominance relation for any subset of criteria P . The fuzzy P dominance relation $FD_P : U \times U \rightarrow [0, 1]$ is defined as

$$FD_P(x, y) = \otimes_{i \in P} (f_i(x) \succsim_i f_i(y)),$$

if $x \neq y$ and $FD_P(x, x) = 1$ for $x \in U$.

The P -gradual dominance principle is then defined for each $x, y \in U$ as $FD_P(x, y) \leq FD_d(x, y)$, and an object x is P -consistent in terms of FDRSA if, for any $y \in U$, both (x, y) and (y, x) satisfy the P -gradual dominance principle. The principle means that if x is more dominant than y in the evaluations of P , then x is given a higher class assignment than y . The set of all P -consistent objects in the POU DT T_f is denoted by $GDCON_P(T_f)$. As above, a criterion $i \in P$ is dispensable if $GDCON_{P-\{i\}}(T_f) = GDCON_P(T_f)$. Every minimal subset $P \subseteq C$ such that $GDCON_P(T_f) = GDCON_C(T_f)$ is a reduct of T_f , and the core of T_f is still the set of all indispensable attributes in C .

Let θ_{\succsim} denote $(\succsim_i)_{i \in A}$. Then, by definition, $\mathfrak{F}_{\theta_{\succsim}}(P, d)$ contains all pairs (x, y) that satisfy $\otimes_{i \in P} (f_i(x) \succsim_i f_i(y)) \leq (f_d(x) \succsim_i f_d(y))$. However, according to the definition of \succsim_i , (x, x) does not necessarily satisfy the $\mathfrak{F}_{\theta_{\succsim}}(P, d)$ -dependency constraint, but it trivially satisfies the P -gradual dominance principle. To circumvent this incompatibility, we slightly extend $\mathfrak{F}_{\theta_{\succsim}}(P, d)$ with these trivial pairs. Thus, we define

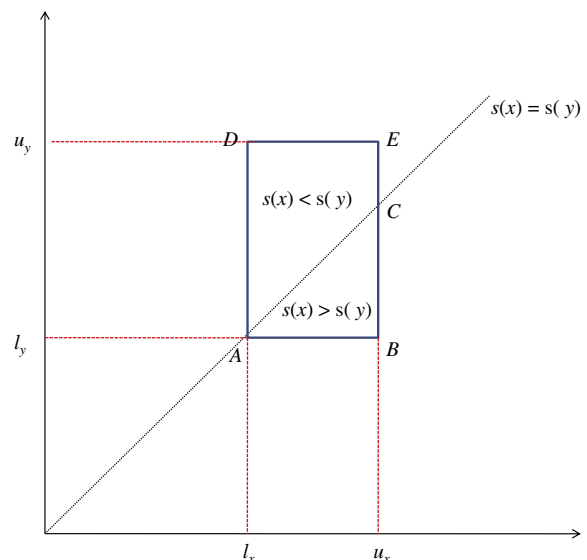


Fig. 1. The degree of dominance between x and y .

$$\mathfrak{F}_{\theta_c}^+(P, d) = \mathfrak{F}_{\theta_c}(P, d) \cup \{(x, x) | x \in U\}.$$

Then, the following lemma and theorem can be derived:

Lemma 5. For any subset of criteria P and any $x, y \in U$,

1. (x, y) satisfies the P -gradual dominance principle iff $(x, y) \in \mathfrak{F}_{\theta_c}^+(P, d)$ and
2. $x \in (0\mathfrak{F}_{\theta_c}^+(P, d))^{\text{tr}}$ iff $x \in \text{GD CON}_P(T_f)$

Theorem 5. A subset of criteria P is a reduct (resp. the core) of T_f in FDRSA iff it is an object-based reduct (resp. the object-based core) of T_f with respect to the $\mathfrak{F}_{\theta_c}^+(C, d)$ dependency constraint.

Example 7. Continuing with Example 3, according to the definition of \sim_i , we can construct the following matrix, whose (x, y) -entry denotes $(f_1(x) \sim_1 f_1(y), f_2(x) \sim_2 f_2(y), f_d(x) \sim_d f_d(y))$.

	1	2	3	4	5	6	7	8
1	-	$(\frac{7}{8}, \frac{3}{4}, 1)$	$(\frac{3}{7}, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$
2	$(\frac{1}{8}, \frac{1}{4}, 1)$	-	$(\frac{9}{56}, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$
3	$(\frac{4}{7}, 0, 0)^*$	$(\frac{47}{56}, 0, 0)^*$	-	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$	$(1, 1, 1)$
4	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 1)$	-	$(0, \frac{19}{21}, 1)$	$(\frac{2}{3}, \frac{2}{7}, 1)^*$	$(1, 1, 1)$	$(1, 1, 1)$
5	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 0)$	$(0, \frac{2}{21}, 0)^{**}$	-	$(1, 0, 0)^*$	<u>$(1, 1, 0)$</u>	<u>$(1, 1, 0)$</u>
6	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 0)$	<u>$(\frac{1}{3}, \frac{5}{7}, 0)$</u>	$(0, 1, 1)$	-	<u>$(1, 1, \frac{1}{2})$</u>	<u>$(1, 1, \frac{1}{2})$</u>
7	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 1)$	$(0, 0, \frac{1}{2})$	-	$(0, 1, \frac{1}{2})^{**}$
8	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 0)$	$(0, 0, 1)$	$(0, 0, \frac{1}{2})$	$(1, 0, \frac{1}{2})^*$	-

Note that we leave all entries (x, x) blank because all such pairs trivially belong to the (extended) dependency constraint; therefore, we do not bother to calculate their fuzzy dominance values. Let us take a closer look at the calculation process of some entries in the matrix. For example, $f_1(6) = [24, 27]$ and $f_1(4) = 26$ imply that

$$(f_1(4) \sim_1 f_1(6)) = \frac{\mu_i(\{26\} \times [24, 26])}{\mu_i(\{26\} \times [24, 27])} = \frac{26 - 24}{27 - 24} = \frac{2}{3},$$

so the first component of the tuple in $(4, 6)$ is $\frac{2}{3}$. Another example is $(f_2(2) \sim_2 f_2(1))$ which is equal to $[48, 50] \sim_2 [48, 52]$ and is an instance of Fig. 1, where $l_y = 48$, $u_x = 50$, and $u_y = 52$. Thus, $(f_2(2) \sim_2 f_2(1)) = \frac{50 - 48}{2(52 - 48)} = \frac{1}{4}$, which is exactly the second component of the tuple in $(2, 1)$.

Now, let C be the set of all condition attributes and let \otimes be the Łukasiewicz t-norm, i.e., $a \otimes b = \max(x + y - 1, 0)$. Then, it is clear that the only pairs violating the dependency constraint are $(5, 7)$, $(5, 8)$, $(6, 4)$, $(6, 7)$, $(6, 8)$ (the underlined entries in the matrix), which are also exactly the pairs that violate the C -gradual dominance principle. Thus, we verify Lemma 5(1). In addition, it is easy to verify that $\mathfrak{F}_{\theta_c}^+(C, d)^{\text{tr}} = \text{GD CON}_C(T_f) = \{1, 2, 3\}$, as is expected by Lemma 5(2). Let $P = \{1\}$ and $P' = \{2\}$. Then, in addition to the underlined entries, the $*$ -entries $(3, 1)$, $(3, 2)$, $(5, 6)$, $(8, 7)$ violate the $\mathfrak{F}_{\theta_c}^+(P, d)$ constraint and the $**$ -entries $(5, 4)$, $(6, 4)$ violate the $\mathfrak{F}_{\theta_c}^+(P', d)$ constraint. Thus, the unique relation-based reduct of the POUDT is C . Furthermore, $\mathfrak{F}_{\theta_c}^+(P, d)^{\text{tr}} = \text{GD CON}_P(T_f) = \emptyset$ and $\mathfrak{F}_{\theta_c}^+(P', d)^{\text{tr}} = \text{GD CON}_{P'}(T_f) = \{1, 2, 3\}$, so P is a unique object-based reduct, which is also exactly the reduct according to FDRSA.

5. Conclusion

We have presented an abstract framework that defines object-based and relation-based reducts for attribute reduction in rough set theory. The framework unifies different rough set approaches for attribute reduction in various kinds of information systems from a relational perspective. While the object-based reduct ensures that consistent objects remain unchanged during the elimination of dispensable attributes, the relation-based reduct requires that the pairs of objects satisfying some particular constraint must be retained.

The value of our framework is twofold. On one hand, the *uniformity* of the framework explicates the common principle behind a variety of rough set approaches for attribute reduction and highlights their differences. The common principle indicates that attribute reduction amounts to finding reducts that preserve certain kinds of *dependency constraints* between condition and decision attributes. These dependency constraints are characterized by the relational comparators between the attribute-values of objects,

the operations used to aggregate the results of comparisons, and the implication functions of the aggregated results. By changing the three parameters, we can instantiate the general framework to a particular rough set approach for attribute reduction in a special kind of information system. Thus, we have common ground to compare the differences between these approaches.

On the other hand, the *generality* of the framework extends the application scope of rough set approaches. In our case studies, we investigate the notions of reduct and core in different FIS, so only the object-based reduct and core are considered. In other words, the RIS $\mathcal{O}(T_f)$ only plays an auxiliary role in the study of an FIS T_f . Nevertheless, when we consider a primitive RIS, say, the representation of a social network, we may be interested in the discovery of rules in the form $\bigwedge_{i \in B} r_i(x, y) \rightarrow r_d(x, y)$. In such cases, it is more appropriate to consider attribute reduction with the relation-based reduct and core. While this is a theoretical framework, its practical implication is that many existing techniques, such as the discernibility matrix approach, can be easily adapted to new applications.

In addition, the current definition of RIS can only be applied to the representation of binary relations. However, for complex networks, we may have to represent high-dimensional relations. Because of the generality of our framework, it seems quite feasible that it can be used to represent such complex relational networks and thereby facilitate knowledge reduction in network mining tasks. Let us be more specific on this point and express it in a formal logical language. All dependency constraints in this paper can be expressed as a first-order logic formula with two

free variables⁶ $\varphi_B(x,y)$, where B is the set of condition attributes in the formula. For example, the indiscernibility constraint can be written as

$$\varphi_B(x,y) = \bigwedge_{i \in B} f_i(x) = f_i(y) \rightarrow d(x) = d(y),$$

where each f_i and d are function symbols corresponding to the attributes. Now, let us consider a complex social network. A k -ary relation among the individuals in such a network can be represented as a k -place predicate symbol in first-order logic. Thus, various dependency constraints in a complex social network can be expressed as first-order logic formulas with a set of predicate symbols of arbitrary arity. Consequently, we can extend the application scope of rough set approaches from attribute reduction in functional information systems to knowledge reduction in complex relational networks.

Acknowledgments

This work was partially supported by the National Science Council of Taiwan under grants 98-2221-E-001-013-MY3 and NSC 99-2410-H-346-001. We wish to thank the editor and three anonymous referees for their constructive suggestions.

References

- Dembczynski, K., Greco, S., Slowinski, R., 2009. Rough set approach to multiple criteria classification with imprecise evaluations and assignments. *European Journal of Operational Research* 198 (2), 626–636.
- Fan, T., Liu, D., Tzeng, G., 2006. Arrow decision logic for relational information systems. *LNCSTransactions on Rough Sets V (LNCST 4100)*, pp. 240–262.
- Fortemps, P., Greco, S., Slowinski, R., 2008. Multicriteria decision support using rules that represent rough-graded preference relations. *European Journal of Operational Research* 188 (1), 206–223.
- Greco, S., Matarazzo, B., Slowinski, R., 1999. Rough approximation of a preference relation by dominance relations. *European Journal of Operational Research* 117 (1), 63–83.
- Greco, S., Matarazzo, B., Slowinski, R., 2001. Rough set theory for multicriteria decision analysis. *European Journal of Operational Research* 129 (1), 1–47.
- Greco, S., Matarazzo, B., Slowinski, R., 2008. Fuzzy set extensions of the dominance-based rough set approach. In: Bustince, H., Herrera, F., Montero, J. (Eds.), *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*. Springer-Verlag, pp. 239–261.
- Greco, S., Matarazzo, B., Slowinski, R., 2010. Dominance-based rough set approach to decision under uncertainty and time preference. *Annals of Operations Research* 176 (1), 41–75.
- Hájek, P., 1998. *Metamathematics of Fuzzy Logic*. Kluwer Academic Publisher.
- Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. *Bulletin de la Societe de Vaud des Sciences Naturelles* 44, 223.
- Kryszkiewicz, M., 1998. Properties of incomplete information systems in the framework of rough sets. In: Polkowski, L., Skowron, A. (Eds.), *Rough Sets in Knowledge Discovery*. Physica-Verlag, pp. 422–450.
- Kryszkiewicz, M., Rybiński, H., 1996a. Reducing information systems with uncertain attributes. In: Raś, Z.W., Michalewicz, M. (Eds.), *Proceedings of the Ninth ISMIS*. LNAI 1079. Springer-Verlag, pp. 285–294.
- Kryszkiewicz, M., Rybiński, H., 1996b. Reducing information systems with uncertain real value attributes. In: *Proceedings of the Sixth International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems*. pp. 1165–1169.
- Liau, C., Lin, T., 2005. Reasoning about relational granulation in modal logics. In: *Proceedings of the First IEEE International Conference on Granular Computing*. IEEE Press, pp. 534–538.
- Mendelson, E., 1997. *Introduction to Mathematical Logic*, forth ed. Chapman & Hall/CRC.
- Pawlak, Z., 1982. Rough sets. *International Journal of Computer and Information Sciences* 11 (15), 341–356.
- Pawlak, Z., 1991. *Rough Sets – Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers..
- Skowron, A., Rauszer, C., 1991. The discernibility matrices and functions in information systems. In: Slowinski, R. (Ed.), *Intelligent Decision Support Systems: Handbook of Applications and Advances in Rough Set Theory*. Kluwer Academic Publisher, pp. 331–362.
- Skowron, A., Stepaniuk, J., 1996. Tolerance approximation spaces. *Fundamenta Informaticae* 27 (2–3), 245–253.
- Slowinski, R., Greco, S., Matarazzo, B., 2002. Rough set analysis of preference-ordered data. In: Alpigini, J., Peters, J., Skowron, A., Zhong, N. (Eds.), *Proceedings of the Third International Conference on Rough Sets and Current Trends in Computing*. LNAI 2475. Springer-Verlag, pp. 44–59.
- Slowinski, R., Greco, S., Matarazzo, B., 2009. Rough sets in decision making. In: Meyers, R.A. (Ed.), *Encyclopedia of Complexity and Systems Science*. Springer, pp. 7753–7786.
- Sokal, R., Sneath, P., 1963. *Principles of Numerical Taxonomy*. Freeman, San Francisco, CA.
- Sweeney, L., 2002a. Achieving k -anonymity privacy protection using generalization and suppression. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems* 10 (5), 571–588.
- Sweeney, L., 2002b. k -anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems* 10 (5), 557–570.
- Wang, D., Liau, C., Hsu, T., 2004. Medical privacy protection based on granular computing. *Artificial Intelligence in Medicine* 32 (2), 137–149.
- Wang, D., Liau, C., Hsu, T., 2007. An epistemic framework for privacy protection in database linking. *Data and Knowledge Engineering* 61 (1), 176–205.
- Wang, D., Liau, C., Hsu, T., Chen, J., 2006. Value and damage of information: A data privacy perspective. *International Journal of Approximate Reasoning*.

⁶ See Mendelson (1997) for an elementary introduction to first-order logic.