ELSEVIER

Information Sciences 124 (2000) 301–316

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

α-RST: a generalization of rough set theory

Mohamed Quafafou

IRIN, University of Nantes, 2 rue de la Houssiniere, BP 92208-44322, Nantes Cedex 03, France

Received 16 September 1997; received in revised form 13 July 1999; accepted 4 September 1999

Abstract

The paper presents a transition from the crisp rough set theory to a fuzzy one, called Alpha Rough Set Theory or, in short, α-RST. All basic concepts or rough set theory are extended, i.e., information system, indiscernibility, dependency, reduction, core, definability, approximations and boundary. The resulted theory takes into account fuzzy data and allows the approximation of fuzzy concepts. Besides, the control of knowledge granularity is natural in α-RST which is based on a parameterized indiscernibility relation. α-RST is developed to recognize non-deterministic relationships using notions as α-dependency, α-reduct and so forth. On the other hand, we introduce a notion of relative dependency as an alternative of the absolute definibility presented in rough set theory. The extension α-RST leads naturally to the new concept of alpha rough sets which represents sets with fuzzy non-empty boundaries.

Keywords: Rough sets; Fuzzy sets; Attributes dependency; Concept approximation

1. Introduction

Rough sets are a suitable mathematical model of vague concepts, i.e., concepts without sharp boundaries. Rough set theory is emerging as a powerful theory dealing with imperfect data [1–3]. It is an expending research area which stimulates explorations on both real-world applications and on the theory itself, i.e., decision analysis, machine learning, knowledge discovery, market research, conflict analysis, and so forth. Recent theoretical developments on this theory and its applications are collected in [4].

E-mail address: quafafou@irin.univ-nantes.fr (M. Quafafou).

0020-0255/00/$ - see front matter
PII: S 0 0 2 0 - 0 2 5 5 (9 9) 0 0 0 7 5 - 4

Rough set theory assumes that information systems contain only crisp data and any feature (attribute) of any object (example) has a precise and unique value. However, real-world data are generally imprecise, tend to be noisy, contaminated by errors, and values for attributes are often missing. To overcome the problem of uncertain data, Slowinski et al. discuss in Ref. [5] four sources of uncertainty; discretization process [11], imprecise data, missing value and multiple descriptors, and they propose a generalization of the classical indiscernibility relation to handle uncertainty caused by multiple descriptors. On the other hand, Ziarko proposed an extension of rough sets, called *variable precision model of rough set theory* (VPRS), to take into account non-deterministic relationships and to identify strong rules [6,7]. Other extensions to standard rough set theory have been suggested in Refs. [8–10].

In this paper, we present a new extension of rough set theory, called α-RST, where all basic concepts of rough set theory, i.e., information system, indiscernibility relation, dependency, and so forth, are generalized. Fuzzy sets are used to discretize continue attributes but α-RST is developed to approximate fuzzy sets. On the other hand, the control of knowledge granularity is natural in α-RST because it is based on a parameterized indiscernibility relation. Another feature of α-RST is the notion of α-*dependency*, i.e., a set of attributes depends on another with a given degree in the range [0, 1]. This notion may be seen as a partial dependency between attributes. We have also introduced notions as α-*reduction*, α-*definability* and *approximations of fuzzy sets* that are also fuzzy. Finally, α-RST is to allow the control of the universe partitioning and the approximation of concepts.

In this paper we present extensions of the basic concepts of rough set theory. Section 2 presents a generalized information systems. Section 3 introduces the aggregation operator and states the new description of an object of the universe. A parameterized indiscernibility relation based on a similarity threshold is defined in Section 4. The notion of α-core, using extensions of dependency and reduction concepts, is defined in Section 5. Section 6 analyses the notion of definability and proposes a new definition of definability, called α-definability, which may be seen as a relative definability in contrast with the absolute definability in rough set theory. Section 7 is dedicated to approximations of fuzzy concepts and their properties. At the end of this section, the new concept of *Alpha rough sets* immerges, i.e., sets with a fuzzy non-empty boundary. In Section 8, we discuss some important issues through related work and we conclude in Section 9.

2. A generalized information system

Let U be a finite set of objects or examples, $U = \{x_1, x_2, \ldots, x_n\}$. These objects of the universe U are characterized by a finite set of attributes, denoted as $Q = \{q_1, q_2, \ldots, q_m\}$. An attribute q_i may be qualitative or quantitative. A qualitative attribute q_i has a domain v_{qi} which determines the set of possible values for attribute q_i. For instance, the attribute Age may be seen as a qualitative attribute (nominal or categorical) which is defined by its discrete domain (a finite discrete set) as {young, adult, middle_age, old, too_old}. Equivalently, each value of Age is described by its membership function. The membership function of middle_age is depicted in Fig. 1(a). The attribute Age may also be seen as a quantitative (cardinal) one, which is defined by its continuous domain (interval) as $[a, b]$. A preprocessing discretization phase is generally necessary for a numerical attribute to divide its domain into intervals which correspond to qualitative terms. This translation process may be performed automatically [discretization] or by an expert [12]. Fig. 1(b) shows a membership function which represents the value middle_age of attribute Age. Modeling a value of attribute using membership depicted in Fig. 1(a) and (b) assumes that concepts are crisp, i.e., the age of a person is middle_age or is not. This approach is based on 2-valued logic – black or white. Consider next that middle_age is defined by fuzzy property "around 45 years". Qualitative attributes may be interpreted as linguistic variables, according to the fuzzy set theory [13,14]. In this case, there is not a unique membership function for middle_age and the expert has to determine it according to the nature of the application domain. Fig. 1(c) shows the membership function which represent

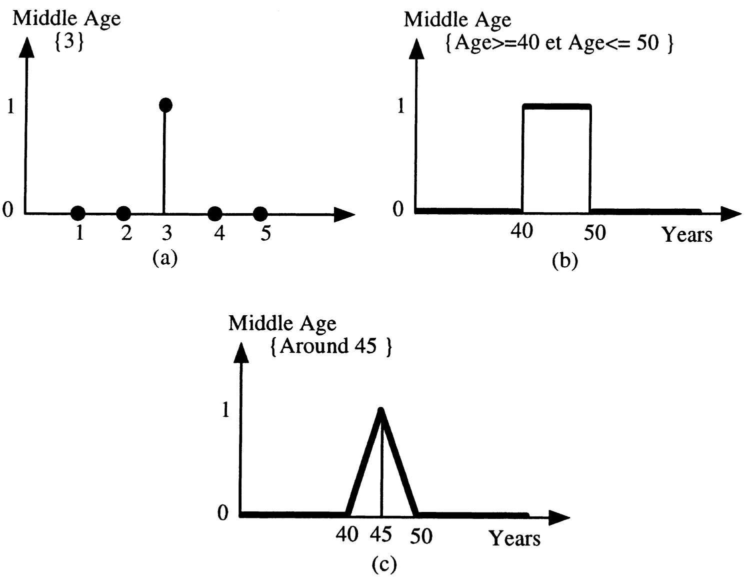

Fig. 1. Possible membership functions to represent midle_age.

the value middle_age; its values measure degrees to which examples satisfy the imprecisely defined property around 45 yr.

Each attribute value is represented by a membership function, which defines a crisp or a fuzzy set. The set V_q is a domain of the attribute q, $V_q = \{v_{q1}, v_{q2}, \ldots, v_{qr}\}$ and V is equal to $\bigcup_{q \in Q} V_q$. We note as M_{qi} the set of membership functions associated to the value v_{qi} of attributes qi, $M_q = \{\mu_{q1}, \mu_{q2}, \ldots, \mu_{qr}\}$ and $M = \bigcup_{q \in Q} M_q$.

So far, we have seen that each value of any attribute is represented with a membership function. This modeling preprocess may influence the results of systems using rough set theory [15,16]. We define the notion of generalized information system in order to take into account uncertainty inherent to both the data and the preprocessing. We define the generalized information function $\tilde{\rho} = \rho \times \delta$, where:

- ρ is called *nominal information function* ρ: $\begin{bmatrix} U \times Q & \to & V, \\ (x,q) & \to & v_{qi}, \end{bmatrix}$
- δ is called *cardinal information function* δ: $\begin{bmatrix} U \times Q & \to & [0,1], \\ (x,q) & \to & \mu_{qi}(q(x)). \end{bmatrix}$

Each object x of the universe U is described by a vector of pairs $(\rho(x,q), \delta(x,q))$ for all q in Q. Thus, $\rho(x,q)$ represents the nominal (linguistic) value of the attribute q for the object x, whereas $\delta(x,q)$ is the cardinal value of q, i.e., degree of possibility that the attribute q has value $\rho(x,q)$ for the object x. A generalized information system represents a function, called a *alpha information function*, denoted as $\tilde{\rho}$, which maps $U \times Q$ into $V \times [0,1]$.

$$\tilde{\rho} : \begin{bmatrix} U \times Q & \to & V \times [0,1], \\ (x,q) & \to & (\rho(x,q), \delta(x,q)). \end{bmatrix}$$

Consequently, we define a generalized information system as a classical information system with nominal and cardinal information functions $\langle U, V, Q, \tilde{\rho} \rangle$.

3. Aggregation operator and fuzzy subsets

We introduce an *aggregation operator*, denoted ϕ_R, which associates to each object x a degree of possibility of its realization according to a set of attributes R. This degree results from the aggregation of $\delta(x,q)$ for all attributes R. In what follows we will define ϕ_R as Yager's parameterized t-norm, with $1 \to \infty$:

$$\phi_R \begin{bmatrix} U & \to & [0,1], \\ x & \to & \min\{\delta(x,q)\ \forall q \in R\}. \end{bmatrix}$$

Consequently, we associate to a given subset of attributes R a region of the universe U, denoted $\tilde{U}_R$ or, $\tilde{U}$ which is defined as the graph of the function ϕ_R, i.e., $\tilde{U} = \text{graph}(\phi_R) = \{(x, \phi_R(x)) : x \in U\}$. $\tilde{U}$ is a fuzzy subset of the universe U, indeed, any object x of U belongs to $\tilde{U}$ to some degree $\phi(x) \in [0, 1]$ and the membership function is not black or white. In what follows, members of $\tilde{U}$ will be noted $\tilde{x}$, with $\tilde{x} = (x,\ \phi_R(x))$.

Proposition 1. *Let* $T = Y_i q_i$ *a subset of* Q, *so* $\text{graph}(\phi_R) = I_i\ \text{graph}(\phi_{qi})$.

The previous proposition holds because $\forall x \in U_R, \phi_R(x) \leqslant \phi_{qi}(x)$ for all $q_i \in T$ and $\phi_R(x) = min_{q_i \in R}(\phi_{qi}(x))$. In this context, the description of an object $x \in U$ in terms of values of attributes from $R \subseteq Q$, denoted $\alpha\text{-Des}_R(x)$ becomes

$$\alpha\text{-Des}_R(x) = \{(q, (v, \pi)) : \tilde{\rho}(x, q) = (v, \pi)\ \forall q \in R\}.$$

Let R be a subset of attributes containing k attributes denoted as $q_1, q_2, \ldots, q_k$. The value of an attribute q_i in a generalized information system is a couple (v_i, π_i), where v_i is a nominal value, i.e., $v_i = \rho(x, q_i)$, and π_i is the degree of possibility for x to have value v_i for attribute q_i, i.e., $\pi_x = \delta(x, q_i)$. Consequently, an object x is characterized by two vectors, denoted vx and πx. The first one $vx = (v_1, v_2, \ldots, v_k)$ contains nominal values, whereas the second one $\pi x = (\pi_1, \pi_2, \ldots, \pi_k)$ contains possibility degrees which a vector in $I^k = [0, 1]^k$. The first vector is the classical description of objects in rough set theory, whereas the second vector is a point in the hypercube I^k, which is the set of all fuzzy subsets. Indeed, a fuzzy set is any point in the cube I^k [17]. The second vector is a fuzzy set.

In the classical rough set theory, an object or an example is represented by a vector in multi-dimensional feature space. This description is known in machine learning as attribute-value description. The indiscernibility relation in rough set theory is defined basing on the equality of attribute values. In contrast, an example is described in α-RST by two vectors resulting from both cardinal and nominal functions. In the next section, we generalize the indiscernibility notion to take into account both the symbolic values and possibility degrees.

4. Generalization of indiscernibility relation

The basic operations in rough set theory are approximations which are defined according to the indiscernibility notion. In fact, objects of the universe may be grouped according to values of a given set of attributes R. Each group contains objects with the same values of attributes in R. This means that the information given by the attributes in R is not sufficient to distinguish objects in the same group. We say that objects of the same group are indiscernible. This equivalence relation, denoted $IND(R)$, is defined on the equality of values of attributes in R. The quotient set $U/IND(R)$ contains equivalence classes which are granules of knowledge representation. These groups, called R-elementary sets, form basic granule of knowledge about the universe and they are used to compute approximations of concepts.

The previous indiscernibility definition is not sufficient when we consider a generalized information system because it does not take into account possibility degrees associated with values of attributes in R. In order to cope with this problem, we introduce a parameterized relation, denoted $IND(R,\alpha)$ or, in short, R_α. We consider that two elements are indiscernible if and only if they have the same values for all attributes and if their possibility degree is greater than a given *similarity threshold*, denoted α:

$$\forall x, y \in U \quad xR_\alpha y \iff x\tilde{R}y \text{ and } \mu_{\tilde{S}}(\pi x, \pi y) \geqslant \alpha.$$

The relation R_α is defined on both an equivalence relation $\tilde{R}$ and a similarity relation $\tilde{S}$. The relation $\tilde{R}$ is similar to the classical indiscernibility relation R, definition is based on the equality of attribute values (i.e. $x\,\tilde{R}\,y \iff vx = vy$). On the other hand, the relation $\tilde{S}$ must verify the three following properties [18]:

1. Reflexivity: $\forall x \in U \quad \mu_{\tilde{S}}(\pi x, \pi x) = 1$,
2. Symmetry: $\forall x, y \in U \quad \mu_{\tilde{S}}(\pi x, \pi y) = \mu_{\tilde{S}}(\pi y, \pi x)$,
3. Max-*transitivity: $\forall x, y, z \in U \quad \mu_{\tilde{S}}(\pi x, \pi z) \geqslant \mu_{\tilde{S}}(\pi x, \pi y) * \mu_{\tilde{S}}(\pi y, \pi z)$,

where $*$ is a binary operation defined on $[0,1]$ such that $0*0 = 1*0 = 0*1 = 0$ and $1*1 = 1$. The operator $*$ is usually chosen such that $\forall (a,b) \in [0,1]^2,\ a * b \leqslant \min(a,b)$. The operator $*$ is assumed to be "min" in what follows. The pioneer works on $*$ are described in Refs. [19–21], see also [22] for more information on similarity relation.

Proposition 2. *If $\tilde{R}$ is an equivalence relation, $\tilde{S}$ is a similarity relation and the operator $*$ is the "min" then R_α is an equivalence relation.*

Proof. Let $\tilde{R}$ relation be an equivalence relation and $\tilde{S}$ be a similarity relation.

1. Reflexivity: $x\tilde{R}x$ and $\mu_{\tilde{S}}(\pi x, \pi x) = 1 \geqslant \alpha$.
2. Symmetry: $xR_\alpha y \Rightarrow \mu_{\tilde{S}}(\pi x, \pi y) \geqslant \alpha$, so $\mu_{\tilde{S}}(\pi y, \pi x) \geqslant \alpha \Rightarrow yR_\alpha x$.
3. Transitivity: $x\ R_\alpha\ y$ and $y\ R_\alpha\ z \Rightarrow \ \mu_{\tilde{S}}(\pi x, \ \pi y) \geqslant \alpha$ and $\mu_{\tilde{S}}(\pi y, \pi z) \geqslant \alpha$, hence, $\mu_{\tilde{S}}(\pi x, \pi z) \geqslant \min(\mu_{\tilde{S}}(\pi x, \pi y), \mu_{\tilde{S}}(\pi y, \pi z)) \geqslant \alpha$.

An equivalence class of U/R_α which is determined by an element $x \in U$ is denoted as $[x]_R^\alpha$. These classes result from the partition of all the universe U. So far, we have associated a fuzzy subset $\widetilde{U}_R = \{(x, \phi_R\ (x)) : x \in U\}$ to each subset R of attributes. The result of the partition of this subset are fuzzy subsets (or classes):

$$\forall \tilde{x}, \tilde{y} \in \tilde{U}_R, \quad \tilde{x} R_\alpha \tilde{y} \quad \Longleftrightarrow \quad \tilde{x}\ \tilde{R}\ \tilde{y} \text{ and } \mu_{\tilde{S}}(\pi\tilde{x}, \pi\tilde{y}) \geqslant \alpha.$$

The family of all equivalence classes of relation R_α on $\tilde{U}_R$ is denoted by $\tilde{U}/R_\alpha$. The following proposition states the relation between U/R_α and $\tilde{U}/R_\alpha$. The family of all equivalence classes of relation R_α on $\tilde{U}_R$ is denoted by $\tilde{U}/R_\alpha$.

The indiscernibility relation defined in rough set theory is black or white, i.e., two elements are indiscernible or they are not. The alpha indiscernibility, denoted α-IND, generalizes the notion of indiscernibility by computing an indiscernibility degree associated to any two objects. The degree of indiscernibility of two objects $\tilde{x}$ and $\tilde{y}$, denoted α-IND $(\tilde{x}, \tilde{y})$, corresponds to the degree of non-similarity of the two objects, i.e., α-IND $(\tilde{x}, \tilde{y}) = 1 - \mu_{\tilde{S}}(\tilde{x}, \tilde{y})$. So, elements are more indiscernible when they are more similar.

5. Dependency and reduction

In rough set theory, we say that a set of attributes R *depends* on a set of attributes P, denoted $P \rightarrow R$, iff all elementary sets of the indiscernibility relation associated with P are subsets of some elementary sets defined by R. A set attributes may be reduced in such a way that the resulting set of attributes, which is smaller, provides the same quality of classification as the original sset of attributes. This means that elementary sets generated by the reduced set A are identical to those generated by the original set of attributes P, i.e., $\mathrm{IND}(A) = \mathrm{IND}(P)$. The smallest reduced set of P is called a *reduct*. In the case where an information system has more than one reduct, the intersection of these reducts is computed. It is called the *core* and represents the most significant attributes in the system. The following proposition states the relation between the dependency and reduction notions.

Proposition 3. *Let P and R be two subsets of attributes, $R \subseteq P \subseteq Q$. R is a reduct of P iff $(P \rightarrow R)$ and $(R \rightarrow P)$.*

We extend this notion of dependency to consider generalized information systems. Thus, the dependency relation is not a black or white relation. We say that a set of attributes R depends on a set P iff each elementary set X of indiscernibility relation associated with P has a non-empty intersection with at least an elementary set Y defined by R, and the inclusion degree of X in Y is greater than a dependency parameter, noted as β. We call this property *Alpha dependency* of attributes.

Definition 1 (α-Dependency). Let P and R be two subsets of attributes, $R \subseteq P \subseteq Q$ and $\alpha \in [0, 1]$. R alpha-depends on P iff $\exists\ \beta \in [0, 1]$ such that

$$P \stackrel{\beta}{\rightarrow} R \iff \forall B \in U/IND(P,\alpha),\ \exists b' \in U/IND(R,\alpha)\ \text{degree}(B \subseteq b') \geqslant \beta.$$

The previous definition introduces the notion of *alpha dependency* which can be seen as a partial dependency between attributes. Consequently, the values of attributes in R are partially determined by values of attributes in P. We say that R partially explains P and there is only a partial functional dependency between values of R and P.

Proposition 4. *Let P and R be two subsets of attributes, $\alpha\beta \in [0, 1]$ such that $\alpha \leqslant \beta$, we have $(P \stackrel{\beta}{\rightarrow} R) \Rightarrow (P \stackrel{\alpha}{\rightarrow} R)$.*

Now, we introduce the notion of α-reduct by generalizing Proposition 4.

Definition 2 (α-Reduct). Let P and R be two subsets of attributes, such that $R \subseteq P \subseteq Q$. R is an alpha-reduct of P, i.e., $R = \alpha$-reduct(P), iff $\exists \beta \in [0, 1]$ such that (i) $P \stackrel{\beta}{\rightarrow} R$, $R \stackrel{\beta}{\rightarrow} P$ and (ii) R is minimal.

R is minimal means that there is no subset of attributes $T \subseteq P$ such that $T \stackrel{\alpha'}{\rightarrow} R$ and $\alpha' \geqslant \alpha$. As a generalized information system may have more than one Alpha reduct, we generalize the notion of the core of an attribute P by introducing the notion of Alpha core. The α-core is then defined as the intersection of all α- reducts.

In this section, we have generalized two key concepts of the rough set theory, which are dependency and reduction. The notion of the core is then generalized based on the definition of alpha reduct. These concepts are generally used in the rough set analysis to construct minimal subsets of attributes (reducts) which have the same indiscernibility power as the whole set of attributes. This analysis leads to the construction of deterministic reductions and a deterministic core, i.e., the decision depends on reduction and the core is an essential part of the whole set of attributes. Alpha rough set allows the user to explore the information using different thresholds related to the reduction or the core. This analysis may lead to the construction of strong reduction and cores, which are only consistent with a part of the data in the information system and we may regard the remaining inconsistency information as noisy data.

6. On the definibility of sets

The notion of definability is important because it is at the basis of the rough set definition. Indeed, Pawlak defines a rough set as a set with a non-empty boundary region, which means that its lower approximation is not equal to its upper approximation [1]. Hence, a subset X of the universe U is definable if and only if $\underline{R}X = \bar{R}X$,otherwise X is undefinable. This definition of definability is suitable when all data in the information system are crisp. We have extended it to consider imprecision inherent to real world data, so we can say that a set, or more precisely a fuzzy set, X is R-definable with a given degree α. For instance, X can be R-definable ($\alpha = 1$), strongly R-definable (e.g., $\alpha = 0.9$), slightly R-definable (e.g., $\alpha = 0.25$), and so forth. We call this property the *α-definability* of a subset, and the degree of definability of a set X will be denoted as α-def(X). As a rough set is characterized by its lower and upper approximation, it is natural to consider that the degree of definability of a set depends on its lower and upper approximation. The alpha definability notion is formalized as follows:

$$\alpha - \mathrm{def}(X) = f(\underline{R}_\alpha X, \bar{R}_\alpha X),$$

where

$$\begin{cases} f(\underline{R}_\alpha X, \bar{R}_\alpha X) = 0 & \text{if } \underline{R}_\alpha X = \emptyset, \\ f(\underline{R}_\alpha X, \bar{R}_\alpha X) = 1 & \text{if } \underline{R}_\alpha X = \bar{R}_\alpha X, \\ f(\underline{R}_\alpha X, \bar{R}_\alpha X) \in]0,\ 1[& \text{otherwise.} \end{cases}$$

Any function respecting the previous constraints may be used to compute definability degrees of sets. In what follows we will consider that

$$f(\underline{R}X, \bar{R}X) = \frac{(1+\theta)\ *\ \theta}{2} \quad \text{with} \quad \theta = \mathrm{degree}(\bar{R}X \subseteq \underline{R}X).$$

If a set X is definable, this means that $\underline{R}X = \bar{R}X$, so α-def$(X) = 1$. On the contrary, i.e., $\underline{R}X \neq \bar{R}X$, the definability degree is less than one: α-def$(X) < 1$. Different methods may be used to compute the inclusion degree of a set into another. For instance, let $M(X)$ be the cardinally of the fuzzy set X defined as follows:

$$M(X) = \sum_{e_i \in x} \mu_x(e_i).$$

The degree to which X is a subset of Y is

$$\mathrm{degree}(X \subseteq Y) = \frac{M(X \cup Y)}{M(Y)}.$$

Obviously, any subset X which is the union of elementary sets is 1-definable, i.e., its definability degree is equal to one., Such sets are called definable sets. When a set is not definable we can compute its definability degree and it is said undefinable if the computed degree is none. On the other hand, we define the notion of roughness of a set as a symmetric notion of the definability. Indeed, the roughness of a set as a symmetric notion of the definability. Indeed, the roughness of a set X which is totally definable, i.e. $\alpha = q$, is equal to none. On the other hand, this roughness degree is equal to 1 when X is totally undefinable, i.e. $\alpha = 0$. The roughness of a set X is computed according to its definability degree using the following formula: α-rough $(X) = 1 - \alpha$-def(X). The more sets are undefinable, the more they are rough.

In the rough set theory, a set (concept) can be definable or undefinablel, and the four following classes of undefinable sets are introduced for ranking sets according to their definability property:

B1: if $\underline{R}X \neq \emptyset$ and $\bar{R}X \neq U$, X will be called roughly definable.
B2: if $\underline{R}X \neq \emptyset$ and $\bar{R}X = U$, X will be called externally undefinable.
B3: if $\underline{R}X = \emptyset$ and $\bar{R}X \neq U$, X will be called internally undefinable.
B4: if $\underline{R}X = \emptyset$ and $\bar{R}X = U$, X will be called totally undefinable.

Consequently, if a set X is undefinable and not roughly definable, then it is internally, externally, or totally undefinable. This is a discrete approach for classifying sets according to their definability. However it is not possible to classify sets which belong to the same class, for instance those of the definable sets class or those of the class B_1 (roughly definable sets class). In the α-RST framework, we can compare two sets according to their definability degrees and a concept $C1$ is said to be more definable than concept $C2$ iff α-def (C_1) is greater than α-def(C_2). In conclusion, we note that rough set theory defines an *absolute definability* of a set X by comparing it lower and upper approximation to the empty set and to all the universe, when alpha rough set theory introduces a *relative definability* by comparing its lower approximation to its upper approximation. Any two sets may be classified according to their definability degrees. Finally, Alpha definability is a property of both crisp and fuzzy sets.

7. Approximation of fuzzy concepts

The most important ideas of rough set theory are lower and upper approximations. These approximations are computed for every concept, i.e., the set of all examples for which a value of the decision is fixed. The lower and upper approximations are useful when the information system contains examples which are described by the same value for all condition attributes and belong to different classes, i.e., when the system is inconsistent.

The classical definition of approximation does not hold when we consider generalized information systems, indeed $[\tilde{x}]_{R_\alpha}$ and X are both fuzzy sets. They are extended in order to define approximations of fuzzy sets which are also fuzzy. The basic idea behind degrees of computing process is simple. Let X be a fuzzy subset of U and $x \in U$. The object x belongs to the lower approximation of X iff $[\tilde{x}]_{R_\alpha}$ is included in X, which implies that x belongs to $X \cap [\tilde{x}]_{R\alpha}$. Thus, the degree to which x belongs to the lower approximation is equal to the minimum of degree α_1 to which it belongs to X and degree α_2 to which it belongs to $[\tilde{x}]_{R_\alpha}$. On the other hand, if x belongs to the upper approximation, then x must belong to $[\tilde{x}]_{R_\alpha}$ or to X, thus it belongs to the union of X and $[\tilde{X}]_{R_\alpha}$. In this case the degree of x is equal to the maximum of α_1 and α_2. Consequently, the generalized approximations are defined as follows:

$$\underline{R}_\alpha X = \left\{(x, \underline{\Delta}_x) \in U \times [0,1] \mid [\tilde{x}]_{R_\alpha} \subseteq X \text{ and } \underline{\Delta}_x = \min(\mu_{[\tilde{x}]_{R_\alpha}}(\tilde{x}), \mu_x(\tilde{x}))\right\},$$
$$\bar{R}_\alpha X = \left\{(x, \underline{\Delta}_x) \in U \times [0,1] \mid [\tilde{x}]_{R_\alpha} \cap X \neq \emptyset \text{ and } \underline{\Delta}_x = \max(\mu_{[\tilde{x}]_{R_\alpha}}(\tilde{x}), \mu_x(\tilde{x}))\right\}.$$

Let X and Y be two fuzzy subsets of the universe U, R a subset of attributes, $\alpha \in [0,1]$ a similarity threshold and $\tilde{U}$ a fuzzy subset defined by the graph of the aggregation operator ϕ_R. Lower and upper approximations of fuzzy sets have the following properties.

$\rightarrow \underline{R}_\alpha \emptyset = \emptyset = \underline{R}_\alpha \emptyset$	$\rightarrow \underline{R}_\alpha U \subseteq U$ and $\bar{R}_\alpha U = U$
$\rightarrow \underline{R}_\alpha \tilde{U} = \tilde{U} = \bar{R}_\alpha \tilde{U}$	$\rightarrow \underline{R}_\alpha X \subseteq X \subseteq \bar{R}_\alpha X$
$\rightarrow \underline{R}_\alpha X \cup \underline{R}_\alpha Y \subseteq \underline{R}_\alpha (X \cup Y)$	$\rightarrow \bar{R}_\alpha X \cup \bar{R}_\alpha Y = \bar{R}_\alpha (X \cup Y)$
$\rightarrow R_\alpha (x \cap Y) \subseteq \bar{R}_\alpha X \cap \bar{R}_\alpha Y$	$\rightarrow \underline{R}_\alpha (X - Y) \subseteq \underline{R}_\alpha X - \underline{R}_\alpha Y$
$\rightarrow \bar{R}_\alpha X - \bar{R}_\alpha Y \subseteq \bar{R}_\alpha (X - Y)$	$\rightarrow \underline{R}_\alpha (\underline{R}_\alpha X) = \underline{R}_\alpha X$
$\rightarrow \underline{R}_\alpha X \subseteq \bar{R}_\alpha (\underline{R}_\alpha X)$	$\rightarrow \underline{R}_\alpha (\bar{R}_\alpha X \subseteq \bar{R}_\alpha X$
$\rightarrow \bar{R}_\alpha X = \bar{R}_\alpha (\bar{R}_\alpha X)$	

One can see that the universe U is not definable, whereas the fuzzy subset $\tilde{U}$ is definable. All properties of classical approximations do not hold when we consider approximation of fuzzy sets because the intersection of a fuzzy set and its complement may be non-empty, and the union of fuzzy set and its complement is not equal to the universe U. Lower and upper approximations are at the basis of definition of the boundary of a set. Let R be a subset of attributes, X be a fuzzy subset and $\alpha \in [0,1]$ be a similarity threshold. The lower approximation of X, i.e., $\underline{R}_\alpha X$, and the upper approximation of X, i.e., $\bar{R}_\alpha X$, are also fuzzy sets. We define the boundary of X, denoted by $B_R^\alpha(X)$, as defined in rough set theory

$$B_R^\alpha(X) = \bar{R}_\alpha X,$$

For instance, the concept of *beautiful sight* is vague, i.e., concept without a sharp boundary, indeed, there are some sights for which we cannot decide if they are beautiful or not. This classical approach of rough set is governed by a logic that permits a vague concept to possess one of only two values which are true or false. In Alpha rough set, the boundary of a fuzzy set is also a fuzzy set and an example has a degree of membership in a boundary set A set with a non-empty boundary is not only rough but it is also fuzzy. The fuzziness of the boundary results from the fact that concept are not simply true or false, but may be true to any degree.

8. Discussion through related work

The goal of this first presentation of α-RST is to introduce the main concepts which generalize the basic concepts of rough set theory. The generalization proposed here results from our experience and the lessons learnt from the development of a rule induction system based on both fuzzy and rough sets [23,24]. In what follows we sketch some important issues through related work.

Information systems and indiscernibility relation. Slowinski et al. have proposed a generalized system in [25] considering the case where the value of an attribute may be a set of qualitative terms. They generalize the indiscernibility notion in order to deal with relative degree of possibility of sub-object. This approach does not consider approximation of fuzzy sets even if numerical values are replaced by fuzzy intervals in the discretization process of numerical attributes. Another work [26] was also developed by Slowinski et al. to transform the classical indiscernibility relation used in rough set theory to a more general similarity relation. Dubios and Prade have introduced the concept of twofold fuzzy sets in Ref. [27] and they have also proposed to couple fuzzy sets (vagueness) and rough sets (coarseness) to get a more accurate account of imperfect data [28,29].

Control of knowledge granularity. In rough set theory, any subset of attribute R generates an indiscernibility relation IND(R). Equivalence classes correspond to granules of knowledge representation, which are at the basis of definitions of key concepts, i.e., approximation, dependency and reduction. Ziarko and al. have proposed to control the degree of granularity using the DataLogic system [30]. The control of granularity is now possible using only extension of the basic concepts of rough set theory (α-RST framework), indeed, we have defined a parameterized indiscernibility relation R_α, where the similarity threshold α is a parameter in the range $[0, 1]$. The user can control the partitioning of the universe by varying α from none (coarsest partitioning) to one (finest partitioning).

Partial dependency and reduction of attributes. Dependency and reduction are two important issues in rough sets, indeed discovering dependencies among attributes leads to reduction which are minimal subsets which have the same quality of classification as the original set of attributes. Discovering dependencies is primarily of importance in rough set approach [31] to knowledge analysis, data exploration, learning and more generally reasoning on data. However, real world data may be noisy, contaminated by errors and/or it may not be sufficient to allow the identification of a functional dependency between two subsets of attributes. To deal with this problem, a partial dependency (α-dependency) and a partial core (α-core) are introduced. Partial vision leads to the construction of strong reductions and cores, which are only consistent with a part of the data in the information system. This approach is more robust in dealing with noisy data and consequently enhances the applicability if rough set theory to real-world problems.

Unfortunately it is not possible to discuss all important issues of α-RST in this paper, i.e., approximation of fuzzy concepts, learning strong rules, decision analysis with quantitative information, conditions under which α-RST behave as rough set theory, learning strong rules, vagueness and uncertainty, and so forth. Other works are necessary to clarify the real role of α-RST considering different application areas. We have developed a first learning prototype, called *Alpha*, which is based on local covering notion and the learning approach developed by Grzymala-Busse and Chan [32–34]. Using Alpha prototype, one can have different views of the data by varying the different parameters. Indeed, these latter allow the user to influence the learning process and to explore the data at different levels of abstraction according to a fine or a coarse partitioning of the universe. Finally, rules may be learnt only when they meet some user's criteria.

9. Conclusion

In this paper a new extension of rough set theory, called α-RST, has been presented. α-RSt offers a suitable framework for dealing with uncertain data and for approximation of fuzzy concepts. It allows the control of knowledge granularity and takes into account non-deterministic relationships. The user can have different views of its data by varying different parameters in the range $[0, 1]$. We continue the development of α-RST by studying the relationship between our approach and related work developed by both rough sets and fuzzy set communities. We are also developing different applications on learning and knowledge discovery using α-RST.

Acknowledgements

The author gratefully acknowledges long discussions with Abdesslam Boulkhemair who has helped to improve this paper significantly.

Appendix A

$$\rightarrow \bar{R}_\alpha(X - Y) \supseteq \bar{R}_\alpha X - \bar{R}_\alpha Y$$

$$\begin{aligned}
\mu_{\bar{R}_\alpha(X-Y)}(x) &= \mu_{\bar{R}_\alpha(X\cap Y^c)}(x)\\
\mu_{\bar{R}_\alpha(X\cap Y^c)}(x) &= \max(\phi(x), \mu_{X\cap Y^c}(x))\\
&= \max(\phi(x), \min(\mu_X(x), \mu_x(x), \mu_{Y^c}(x)))\\
&= \max(\phi(x), \min(\mu_X(x), 1-\mu_Y(x)))
\end{aligned}$$

$$\begin{aligned}
\mu_{\bar{R}\alpha X-\bar{R}\alpha Y}(x) &= \mu_{\bar{R}_\alpha X\cap(\bar{R}_\alpha Y)^c}(x)\\
&= \min(\mu_{\bar{R}_\alpha X}(x), \mu_{(\bar{R}_\alpha Y)^c}(x))\\
&= \min(\mu_{\bar{R}_\alpha X}(x), 1-\mu_{\bar{R}_\alpha Y}(x))\\
&= \min(\max(\phi(x), \mu_X(x)), 1-\max(\phi(x), \mu_Y(x)))
\end{aligned}$$

$$\begin{aligned}
\mu_{\bar{R}_{\alpha(X\cap Y^c)}}(x) &= \max(a, \min(b, 1-c))\\
\mu_{\bar{R}_\alpha X-\bar{R}_\alpha Y}(x) &= \min(\max(a,b), 1-\max(a,c))
\end{aligned}$$

we have to prove that

$$\max(a, \min(b, 1-c)) \geqslant \min(\max(a,b), 1-\max(a,c)) \quad \text{for all } a, b, c \in [0,1]$$

note that

$$1-\max(a,b) = 1-\min(-a,-c) = \min(1-a, 1-c) \tag{A.1}$$

and

$$\max(a+x, b+y) \leqslant \max(a,b) + \max(x,y) \tag{A.2}$$

(i) If $\max(a,b) \leqslant 1-\max(a,c)$: Using (1), we obtain $1-\max(a,c) = \min(1-a, 1-c)$. So,

$$\max(a,b) \leqslant \min(1-a, 1-c) \quad \text{and} \quad b \leqslant 1-c. \tag{A.3}$$

(A.3) $\Rightarrow \min(b, 1-c) = b$ and $b \leqslant \max(a, \min(b, 1-c))$. Similarly, we have $a \leqslant \max(a, \min(b, 1-c))$ and $\max(a,b) \leqslant \max(a, \min(b, 1-c))$.

(ii) If $1-\max(a,c) \leqslant \max(a,b)$,

$$\min(1-a, 1-c) \leqslant \max(a,b) \tag{A.4}$$

(ii.1) If $b \leqslant 1-c \Rightarrow \max(a,b) = \max(a, \min(b, 1-c)$, (4) $\Rightarrow \min(1-a, 1-c) \leqslant \max(a, \min(b, 1-c)$.

(ii.2) If $b > 1-c$, we have $1 \leqslant \max(2a, 1)$ (2) $\Rightarrow \max(2a, 1) \leqslant \max(a,c) + \max(a, 1-c)$ so,

$$1-\max(a,c) \leqslant \max(a, 1-c(\leqslant \max(a, \min(b, 1-c)) \tag{A.5}$$

(A.5) and $\min(b, 1-c) \leqslant b \Rightarrow 1-\max(a,c) \leqslant \max(a,b)$.

References

[1] A. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, Kluwer academic publishers, Dordrecht, The Netherlands, 1991.
[2] A. Pawlak, J.W. Grzymala-Busse, R. Slowinski, W. Ziarko, Rough sets, Communications of the ACM 38 (11) (1995) 89–95.
[3] J.W. Grzymala-Busse, Rough sets, Advances in Imaging and Physics 94 (1995) 151–195.
[4] R. Slowinski, Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, Dordrecht, 1992.
[5] R. Slowinski, J. Stefanowski, Handling various types of uncertainty in the rough set approach, in: Proceedings of the International workshop on Rough Sets and Knowledge Discovery, Springer, Berlin, pp. 366–376.
[6] W. Ziarko, Variable precision rough sets model, Journal of Computer and System Sciences 46 (1) (1993) 39–59.
[7] W. Ziarko, Analysis of uncertain information in the framework of variable precision rough sets, Foundations of Computing and Decision Science 18 (3/4) (1993) 381–396.
[8] J.S. Deogun, Raghavan, H. Sever, Exploiting upper approximation in the rough set methodology, in: International Conference on Knowledge Discovery and Data Mining, 1995, pp. 69–74.
[9] M. Hadjimichael, S.K.K. Wong, Fuzzy representation in rough set approximations, in: International Workshop on rough Sets and Knowledge Discovery in Databases, 1993, pp. 349–356.
[10] Z. Pawlak, S.K.M. Wong,W. Ziarko, Rough Sets: Probabilistic versus Deterministic, in: B.R. Gaines, J.H. Boose (Eds.), Machine Learning and Uncertain Reasoning, Academic Press, New York, 1990, pp. 227–241.
[11] U.M. Fayyad, K.B. Irani, Multi-interval discretization of continuous attributes for classification learning, in: R. Bajcsy (Ed.), Proceedings of the 13th International Joint Conference on Artificial Intelligence, Morgan Kaufmann, Los Altos, CA, 1993.
[12] R. Slowinski, C. Zopounidis, Rough-set sorting of firms according to bankruptcy risk, in: M. Parruccini (Ed.), Applaying Multiple Criteria Aid for Decision to Environmental Management, 1994, pp. 339–357.
[13] R.E. Yager, L.H. Zadeh, An Introduction to Fuzzy Logic Applications in Intelligent Systems, Kluwer Academic Publisher, Norwell, MA, 1992.
[14] E. Cox, The Fuzzy Systems Handbook, Academic Press, New York, 1994.
[15] R. Nowicki, R. Slowinski, J. Stefanowski, Rough set analysis of diagnostic capacity of vabroacoustic symptoms, Computers and Mathematics with Applications 24 (7) (1992) 109–123.
[16] K. Slowinski, R. Slowinski, Sensitivity analysis of rough classification, International Journal of Man–Machine Studies 32 (1990) 693–705.
[17] B. Kosko, Neural Networks and Fuzzy Systems, A Dynamical Systems Approach to Machine Intelligence, Prentice-Hall, Englewood Cliffs, NJ, 1992.
[18] D. Dubois, H. Parde, Similarity-based approximate reasoning, in: Proceedings of the IEEE Symposium Computational Intelligence: Imitating Life, Orlando, FL, 27 June–1 July 1994.
[19] E. Trillas, L. Valverde, A inquirey into indistuinguishability operators, in: H.J. Skala, S. Termini, E. Trillas, D. Reidel (Eds.), Aspects of Vagueness, 1984, pp. 231–256.
[20] J.C. Bezdec, J.D. Harris, Fuzzy partitions and relations: a axiomatic basis for clustering, Fuzzy Sets and Systems 1 (1978) 111–127.
[21] E.H. Ruspini, A theory of fuzzy clustering, in: Proceedings of the IEEE Conference on Decision Control, New Orleans, 1977, pp. 1378–1383.
[22] S.V. Ovchinnikov, Similarity relations fuzzy partitions and fuzzy orderings, Fuzzy Sets and Systems 40 (1991) 107–126.

[23] M. Quafafou, M. Nafia, GAITS: Fuzzy Sets-Based Algorithms for Computing Strategies Using Genetic Algorithms, Fuzzy Logic in Artificial Intelligence FLAI'93, in: Lecture Notes in Artificial Intelligence, Springer, Linz (Austria), 1993.
[24] M. Quafafou, Machine Learning for More Effective and Adaptive Interactions, The East-West International Conferences on Human–Computer Interaction (EWHCI'94), St. Petersburg, Russia, 2–6 August 1994.
[25] R. Slowinski, A generalization of the indiscernability relation for rough set analysis of qualitative information. Rivisita di matematica per le scienze economiche e sociali – Anno 15, Fascicolo 1, 1992, pp 65–78.
[26] R. Slowinski, D. Vanderpooten, Similarity relation as a basis for rough approximations. ISC Research Report 53/95, Warsaw University of Technology, 1995, p. 16.
[27] D. Dubois, H. Prede, Twofold fuzzy sets and rough sets-some issues in knowledge representation, Fuzzy Sets and Systems 23 (1987) 3–18.
[28] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17 (1990) 191–207.
[29] D. Dubois, H. Prade, Putting rough sets and fuzzy sets together, in: Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, Dordrecht, 1991, pp. 203–232.
[30] W. Ziarko, R. Golan, D. Edwards, An application of Datalogic/R knowledge discovery tool to identify strong predictive rules in stock market data, in: Proceedings of AAAI Workshop on Knowledge Discovery in Databases, Washington, DC, 1993, pp. 89–101.
[31] A. Pawlak, R. Slowinski, Decision analysis using rough sets, International Transaction in Operation Research 1 (1) (1994) 107–114.
[32] M.R. Chmielewski, J.W. Grzymala-Busse, N.W. Peterson, S. Than, The rule induction system LERS – a version for personal computers, Foundations of Computing and Decision Sciences 18 (3/4) (1993) 181–212.
[33] C.C. Chan, J.W. Grzymala-Busse, On the two local inductive algorithms: PRISM and LEM2, Foundations of Computing and Decision Sciences 19 (3) (1994) 185–203.
[34] C.C. Chan, J.W. Grzymala-Busse, Rough-Set Boundaries as a tool for learning rules from examples, in: Proceedings of the Fourth Inernational Symposium on Methodology for Intelligent Systems, 12–14 October 1989, pp. 281–288.