

Information Granulation and Rough Set Approximation

Y.Y. Yao

Department of Computer Science, University of Regina
Regina, Saskatchewan, Canada S4S 0A2

E-mail: yyao@cs.uregina.ca

Abstract

Information granulation and concept approximation are some of the fundamental issues of granular computing. Granulation of a universe involves grouping of similar elements into granules to form coarse-grained views of the universe. Approximation of concepts, represented by subsets of the universe, deals with the descriptions of concepts using granules. In the context of rough set theory, this paper examines the two related issues. The granulation structures used by standard rough set theory and the corresponding approximation structures are reviewed. Hierarchical granulation and approximation structures are studied, which results in stratified rough set approximations. A nested sequence of granulations induced by a set of nested equivalence relations leads to a nested sequence of rough set approximations. A multi-level granulation, characterized by a special class of equivalence relations, leads to a more general approximation structure. The notion of neighborhood systems is also explored.

1 Introduction

Granular computing may be regarded to as a label of the family of theories, methodologies, and techniques that make use of granules, i.e., groups, classes, or clusters of a universe, in the process of problem solving [33, 40]. The basic ideas of granular computing have appeared in many fields, such as interval analysis, quantization, rough set theory, Dempster-Shafer theory of belief functions, divide and conquer, cluster analysis, machine learning, databases, information retrieval, and many others [39, 38]. There are many reasons for the study of granular computing [27, 38]. The practical necessity and simplicity in problem solving are perhaps some of the

main reasons. When a problem involves incomplete, uncertain, or vague information, it may be difficult to differentiate distinct elements and one is forced to consider granules. Although detailed information may be available, it may be sufficient to use granules in order to have an efficient and practical solution. Very precise solutions may not be required for many practical problems. The use of granules generally leads to simplification of practical problems. The acquisition of precise information may be too costly, and coarse-grained information reduces cost. There is clearly a need for the systematic studies of granular computing. It is expected that granular computing will play an important role in the design and implementation of efficient and practical intelligent information systems.

The construction, representation, and interpretation of granules, as well as utilization of granules for problem solving, are some of the fundamental issues of granular computing. Information granulation depends on the available knowledge. A granule normally consists of elements that are drawn together by indistinguishability, similarity, proximity, or functionality [30, 38, 39]. An intermediate implication of information granulation is the need for approximation. With the granulated universe, one considers elements within a granule as a whole rather than individually [38]. The loss of information through granulation implies that some subsets of the universe can only be approximately described. We have to deal with approximations of concepts, represented by subsets of the universe, in terms of granules.

A general framework of granular computing was presented in a recent paper by Zadeh [39] in the context of fuzzy set theory. Granules are defined by generalized constraints. Examples of constraints are equality, possibilistic, probabilistic, fuzzy, and veristic constraints. Many specific models of granular computing have also been proposed. Pawlak [16], Polkowski and Skowron [21], and Skowron and Stepaniuk [24] examined granular computing in connection with the theory of rough sets. Yao [32] suggested the use of hierarchical granulations for the study of stratified rough set approximations. Lin [7] and Yao [30, 31] studied granular computing using neighborhood systems. Klir [4] investigated some basic issues of computing with granular probabilities.

Based on these studies, the main objectives of this paper are to investigate the two related issues of information granulation and approximation in the context of rough set theory, and to review studies on these topics. We focus on the analysis of approximation structures with respect to various granulations of the universe. Granulation structures are defined by *similarity* between elements of the universe. The types of similarities range from simple equivalence relations, tolerance relations, and reflexive

binary relations to families of relations, hierarchies, and neighborhood systems. In Section 2, two simple granulation structures are reviewed. The standard rough set theory starts from an equivalence relation. A universe is divided into a family of disjoint subsets. The granulation structure adopted is a partition of the universe, which is well known in mathematics as a quotient set. A pair of lower and upper approximations is used. The approximations are expressed in terms of granules according to their overlaps with the set to be approximated. By weakening the requirement of equivalence relations, we can have more general granulation and approximation structures based on coverings of the universe. In Section 3, hierarchical granulation structures are examined. The notion of stratified rough set approximations is introduced. With respect to different level of granulations, various approximations are obtained. Special types of partition based granulation structures are investigated. A nested sequence of granulations by a nested sequence of equivalence relations leads to a nested sequence of rough set approximations. A hierarchical granulation, characterized by a special class of equivalence relations, leads to a more general approximation structure. For non-partition based granulation structures, we explore the notion of neighborhood systems.

2 Simple Granulations and Approximations

This section reviews some simple granulation structures used in the theory of rough sets. They are characterized by one-level, i.e., single-layered, granulation of the universe.

2.1 Rough set approximations induced by equivalence relations

Let U be a finite and non-empty set called the universe, and let $E \subseteq U \times U$ denote an equivalence relation on U . The pair $apr = (U, E)$ is called an approximation space. The equivalence relation E partitions the set U into disjoint subsets. This partition of the universe is called the quotient set induced by E and is denoted by U/E . The equivalence relation is the available information or knowledge about the objects under consideration. It represents a very special type of similarity between elements of the universe. If two elements x, y in U belong to the same equivalence class, we say that x and y are indistinguishable, i.e., they are similar. Each equivalence class may be viewed as a granule consisting of indistinguishable elements. It is also referred to as

an equivalence granule. The granulation structure induced by an equivalence relation is a partition of the universe.

An arbitrary set $X \subseteq U$ may not necessarily be a union of some equivalence classes. This implies that one may not be able to describe X precisely using the equivalence classes of E . In this case, one may characterize X by a pair of lower and upper approximations:

$$\begin{aligned}\underline{apr}(X) &= \bigcup\{[x]_E \mid x \in U, [x]_E \subseteq X\}, \\ \overline{apr}(X) &= \bigcup\{[x]_E \mid x \in U, [x]_E \cap X \neq \emptyset\},\end{aligned}\quad (1)$$

where

$$[x]_E = \{y \mid y \in U, xEy\}, \quad (2)$$

is the equivalence class containing x . Both lower and upper approximations are unions of some equivalence classes. More precisely, the lower approximation $\underline{apr}(X)$ is the union of those equivalence granules which are subsets of X . The upper approximation $\overline{apr}(X)$ is the union of those equivalence granules which have a non-empty intersection with X .

In addition to the equivalence class oriented definition, i.e., granule oriented definition, we can have an element oriented definition:

$$\begin{aligned}\underline{apr}(X) &= \{x \mid x \in U, [x]_E \subseteq X\}, \\ \overline{apr}(X) &= \{x \mid x \in U, [x]_E \cap X \neq \emptyset\}.\end{aligned}\quad (3)$$

An element $x \in U$ belongs to the lower approximation of X if *all* its equivalent elements belong to X . It belongs to the upper approximation of X if at least *one* of its equivalent elements belongs to X . This interpretation of approximation operators is related to interpretation of the necessity and possibility operators in modal logic [28, 34].

Lower and upper approximations are dual to each other in the sense:

$$\begin{aligned}\text{(Ia)} \quad \underline{apr}(X) &= (\overline{apr}(X^c))^c, \\ \text{(Ib)} \quad \overline{apr}(X) &= (\underline{apr}(X^c))^c,\end{aligned}$$

where $X^c = U - X$ is the complement of X . The set X lies within its lower and upper approximations:

$$\text{(II)} \quad \underline{apr}(X) \subseteq X \subseteq \overline{apr}(X).$$

Intuitively, lower approximation may be understood as the pessimistic view and the upper approximation the optimistic view in approximating a set. One can verify the following properties:

$$\begin{aligned} \text{(IIIa)} \quad & \underline{apr}(X \cap Y) = \underline{apr}(X) \cap \underline{apr}(Y), \\ \text{(IIIb)} \quad & \overline{apr}(X \cup Y) = \overline{apr}(X) \cup \overline{apr}(Y). \end{aligned}$$

The lower approximation of the intersection of a finite number of sets can be obtained from their lower approximations. The similar observation is true for upper approximation of the union of a finite number of sets. However, we only have:

$$\begin{aligned} \text{(IVa)} \quad & \underline{apr}(X \cup Y) \supseteq \underline{apr}(X) \cup \underline{apr}(Y), \\ \text{(IVb)} \quad & \overline{apr}(X \cap Y) \subseteq \overline{apr}(X) \cap \overline{apr}(Y). \end{aligned}$$

One cannot obtain the lower approximation of the union of some sets from their lower approximations, nor obtain the upper approximation of the intersection of some sets from their upper approximations. Additional properties of rough set approximations can be found in Pawlak [14, 15], and Yao and Lin [34].

The accuracy of rough set approximation is defined as [14]:

$$\alpha(X) = \frac{|\underline{apr}(X)|}{|\overline{apr}(X)|}, \quad (4)$$

where $|\cdot|$ denotes the cardinality of a set. For the empty set \emptyset , we define $\alpha(\emptyset) = 1$. Obviously, $0 \leq \alpha(X) \leq 1$. If X is a union of some equivalence granules, we have $\underline{apr}(X) = \overline{apr}(X) = X$, and hence $\alpha(X) = 1$. For $X \neq \emptyset$, $\alpha(X) = 0$ if and only if $\underline{apr}(X) = \emptyset$, independent of its upper approximation. The accuracy measure can be interpreted using the well-known Marczewski-Steinhaus metric, or MZ metric for short. For two sets X and Y , the MZ metric measures the distance between two sets [10]:

$$D(X, Y) = \frac{|X \Delta Y|}{|X \cup Y|} = 1 - \frac{|X \cap Y|}{|X \cup Y|}, \quad (5)$$

where $X \Delta Y = (X \cup Y) - (X \cap Y)$ denotes the symmetric difference between two sets X and Y . It reaches the maximum value of 1 if X and Y are disjoint, i.e., they are totally different, and it reaches the minimum value of 0 if X and Y are exactly the

same. By applying the MZ metric to the lower and upper approximations, we have:

$$\begin{aligned}
 D(\underline{apr}(X), \overline{apr}(X)) &= 1 - \frac{|\underline{apr}(X) \cap \overline{apr}(X)|}{|\underline{apr}(X) \cup \overline{apr}(X)|} \\
 &= 1 - \frac{|\underline{apr}(X)|}{|\overline{apr}(X)|}, \\
 &= 1 - \alpha(X).
 \end{aligned} \tag{6}$$

The accuracy of rough set approximation may be viewed as an *inverse* of MZ metric when applied to lower and upper approximations. In other words, the distance between the lower and upper approximations determines the accuracy of the rough set approximations.

Example 1 The notions of granulation by partitions and rough set approximations can be illustrated by a concrete example using information tables [13]. An information table is a quadruple,

$$T = (U, At, \{V_a \mid a \in At\}, \{f_a \mid a \in At\}),$$

where

- U is a finite and nonempty set of objects,
- At is a finite and nonempty set of attributes,
- V_a is a finite and nonempty set of values for each attribute $a \in At$,
- $f_a : U \longrightarrow V_a$ is an information function for each attribute $a \in At$.

An information table provide a simple, convenient, and powerful tool for describing a set of objects based on their attribute values. Table 1 is an example of information table. We may form granulated views of the universe based on attribute values of objects. For a subset of attributes $A \subseteq At$, we can define an equivalence relation:

$$xE_Ay \iff (\forall a \in A) f_a(x) = f_a(y). \tag{7}$$

Two elements are equivalent (indiscernible) if and only if they have the same value for every attribute in A . The reflexivity, symmetry and transitivity of E_A follow from the properties of the equality relation = between attribute values. For the subset of attributes $\{A_1, A_2\}$, the equivalence relation is defined by the following partition:

$$U/E_{A_1A_2} = \{\{a\}, \{b, c\}, \{d\}, \{e, f\}\}.$$

Object	A_1	A_2	A_3	A_4	Class
a	1	1	1	1	-
b	1	2	1	0	-
c	1	2	0	0	+
d	1	3	1	1	+
e	0	1	0	1	+
f	0	1	1	1	+

Table 1: An information table

Consider the set of objects $+$ = $\{c, d, e, f\}$, we have the rough set approximations:

$$\begin{aligned}\underline{apr}(+) &= \{d\} \cup \{e, f\} = \{d, e, f\}, \\ \overline{apr}(+) &= \{b, c\} \cup \{d\} \cup \{e, f\} = \{b, c, d, e, f\}.\end{aligned}$$

The accuracy of approximation is given by:

$$\alpha(+) = \frac{|\underline{apr}(+)|}{|\overline{apr}(+)|} = \frac{|\{d, e, f\}|}{|\{b, c, d, e, f\}|} = \frac{3}{5}.$$

Similarly, one can choose other subsets of attributes to obtain different approximations of the set $+$. Conceptually, one of the tasks of machine learning, i.e., finding a subset of attributes that properly describe the class $+$, may be formulated as searching for a subset of attributes that produce suitable level of approximations [14]. \square

2.2 Rough set approximations induced by reflexive relations

Let $R \subseteq U \times U$ be a binary relation on U , which is at least reflexive. For two elements $x, y \in U$, if xRy , we say that y is R -related to x . A binary relation may be more conveniently represented using successor neighborhoods, or successor granules [29]:

$$(x)_R = \{y \mid y \in U, xRy\}. \quad (8)$$

The successor neighborhood $(x)_R$ consists of all R -related elements of x . When R is an equivalence relation, $(x)_R$ is the equivalence class containing x . When R is a reflexive relation, the family of successor neighborhoods $U/R = \{(x)_R \mid x \in U\}$ is a covering of the universe, namely, $\bigcup_{x \in U} (x)_R = U$. The binary relation R represents the

similarity between elements of a universe. It is reasonable to assume that similarity is at least reflexive, but not necessarily symmetric and transitive [25].

For the granulation induced by the covering U/R , granule oriented rough set approximations can be defined by generalizing equation (1). The equivalence class $[x]_E$ may be replaced by the successor neighborhood $(x)_R$. One of such generalizations is given by [29]:

$$\begin{aligned}\underline{apr}'(X) &= \bigcup\{(x)_R \mid x \in U, (x)_R \subseteq X\}, \\ \overline{apr}'(X) &= (\underline{apr}'(X^c))^c.\end{aligned}\quad (9)$$

In this definition, we generalize the lower approximation and define the upper approximation through duality. While the lower approximation is the union of some successor neighborhoods, the upper approximation cannot be expressed in this way [29]. The approximations satisfy properties (I), (II), and (IV). They do not satisfy property (III). Nevertheless, they satisfy a weaker version:

$$\begin{aligned}(\text{Va}) \quad & \underline{apr}'(X \cap Y) \subseteq \underline{apr}'(X) \cap \underline{apr}'(Y), \\ (\text{Vb}) \quad & \overline{apr}'(X \cup Y) \supseteq \overline{apr}'(X) \cup \overline{apr}'(Y).\end{aligned}$$

By definition, $\underline{apr}'(X \cap Y)$ can be written as a union of some successor granules. Although both $\underline{apr}'(X)$ and $\underline{apr}'(Y)$ can be expressed as unions of successor granules, $\underline{apr}'(X) \cap \underline{apr}'(Y)$ cannot be so expressed. Alternatively, we generalize the upper approximation and define the lower approximation through duality:

$$\begin{aligned}\underline{apr}''(X) &= (\overline{apr}''(X^c))^c, \\ \overline{apr}''(X) &= \bigcup\{(x)_R \mid x \in U, (x)_R \cap X \neq \emptyset\}.\end{aligned}\quad (10)$$

For a reflexive binary relation, they satisfy properties (I)-(IV). With respect to the element oriented definition, the generalization of equation (3) results in:

$$\begin{aligned}\underline{apr}(X) &= \{x \mid x \in U, (x)_R \subseteq X\}, \\ \overline{apr}(X) &= \{x \mid x \in U, (x)_R \cap X \neq \emptyset\}.\end{aligned}\quad (11)$$

For a reflexive binary relation, they satisfy properties (I)-(IV). In general, these three generalized definitions are not necessarily equivalent to each other. For a reflexive binary relation, they are related to each other by [29]:

$$\underline{apr}''(X) \subseteq \underline{apr}(X) \subseteq \underline{apr}'(X) \subseteq X \subseteq \overline{apr}'(X) \subseteq \overline{apr}(X) \subseteq \overline{apr}''(X). \quad (12)$$

Additional properties of these rough set approximations and their connections can be found in Yao [29].

Example 2 Granulation and approximation in terms of covering can also be illustrated using information tables. A covering of the universe may arise in several ways. One may use binary relations on attribute values to define coverings of the universe [37]. Coverings of universe can also be obtained in set-valued information tables or incomplete information tables [9, 12, 26, 36]. This example uses the latter approach. A set-valued information table is the same as a standard information table except that the information function f_a is a set-valued function, i.e., $f_a : U \longrightarrow 2^{V_a}$. Furthermore, we assume $f_a(x) \neq \emptyset$ for every $a \in At$ and $x \in U$. The set-valued information functions are interpreted as follows. Although an object must take exactly one value from V_a , the available information may be insufficient to determine which value is the actual one. Instead, a set of values is used. Table 2 is a set-valued information table. For a subset of attributes $A \subseteq At$, we can define a compatibility or tolerance relation, i.e., a reflexive and symmetric relation:

$$xS_Ay \iff (\forall a \in A)f_a(x) \cap f_a(y) \neq \emptyset. \quad (13)$$

Two elements are similar if and only if they possibly share the same value for every attribute in A . For the subset of attributes $\{A_1, A_2\}$, the tolerance relation is given by:

$$\begin{aligned} (a)_{S_{A_1A_2}} &= (e)_{S_{A_1A_2}} = \{a, b, d, e\}, \\ (b)_{S_{A_1A_2}} &= (d)_{S_{A_1A_2}} = \{a, b, c, d, e\}, \\ (c)_{S_{A_1A_2}} &= \{b, c, d\}, \\ (f)_{S_{A_1A_2}} &= \{f\}. \end{aligned}$$

They provide a covering of the universe:

$$U/S_{A_1A_2} = \{\{a, b, d, e\}, \{a, b, c, d, e\}, \{b, c, d\}, \{f\}\}.$$

The set of objects $+ = \{c, d, e, f\}$ is approximated by:

$$\begin{aligned} \underline{apr}'(+) &= \underline{apr}''(+) = \underline{apr}(+) = \{f\}, \\ \overline{apr}'(+) &= \overline{apr}''(+) = \overline{apr}(+) = U. \end{aligned}$$

with an accuracy of $1/6$. For an arbitrary set, granule oriented and element oriented approximations may not be the same. For example, for the set $\{b, c, d\}$, we have:

$$\begin{aligned} \underline{apr}'(\{b, c, d\}) &= \{b, c, d\}, & \overline{apr}'(\{b, c, d\}) &= \{a, b, c, d, e\}, \\ \underline{apr}''(\{b, c, d\}) &= \emptyset, & \underline{apr}''(\{b, c, d\}) &= \{a, b, c, d, e\}, \\ \overline{apr}(\{b, c, d\}) &= \{c\}, & \overline{apr}(\{b, c, d\}) &= \{a, b, c, d, e\}. \end{aligned}$$

They clearly satisfy the condition (12). □

Object	A_1	A_2	A_3	A_4	Class
a	$\{0, 1\}$	$\{1\}$	$\{1\}$	$\{0, 1\}$	$\{-\}$
b	$\{1\}$	$\{1, 2\}$	$\{1\}$	$\{0\}$	$\{-\}$
c	$\{1\}$	$\{2\}$	$\{0, 1\}$	$\{0, 1\}$	$\{+\}$
d	$\{1\}$	$\{1, 2, 3\}$	$\{0, 1\}$	$\{1\}$	$\{+\}$
e	$\{0, 1\}$	$\{1\}$	$\{0\}$	$\{0, 1\}$	$\{+\}$
f	$\{0\}$	$\{3\}$	$\{1\}$	$\{1\}$	$\{+\}$

Table 2: A set-valued information table

3 Hierarchical Granulations and Approximations

In the last section, simple one-level granulation structures of the universe are used. The granulated view of the universe is based on a binary relation representing the simplest type of similarities between elements of a universe. Two elements are either related or unrelated. To avoid such a limitation, in this section we examine other types of similarities between objects. More general granulation structures and the corresponding stratified rough set approximations are investigated. Multi-level granulation structures are constructed by putting together simple granulation structures. Each level of the complex structure is a simple granulation structure such as a partition or a covering.

3.1 Nested rough set approximations induced by a nested sequence of equivalence relations

The use of nested sequences of binary relations for defining rough set approximations has been discussed by many authors. Each relation defines a particular type or level of similarities between elements of the universe. Marek and Rasiowa [11] considered gradual approximations of sets based on a descending sequence of equivalence relations. Pomykala [22] used a sequence of tolerance relations. Some recent results on this topic were given by Polkowski [17, 18, 19, 20], Yao [31], and Yao and Lin [35].

A binary relation on U is a subset of the Cartesian product $U \times U$. The set inclusion defines an order on all equivalence relations on U . An equivalence relation E_1 is said to be finer than another equivalence relation E_2 , or E_2 is coarser than E_1 , if $E_1 \subset E_2$. A finer relation produces smaller granules than a coarser relation, i.e., $[x]_{E_1} \subseteq [x]_{E_2}$ for all $x \in U$. Each equivalence granule of E_2 is in fact a union of some

equivalence granules of E_1 . Each granule of E_1 is obtained by further partitioning a granule of E_2 . The relationship between rough set approximations induced by two equivalence relations $E_1 \subseteq E_2$ is given by:

$$E_1 \subseteq E_2 \implies \underline{apr}_{E_2}(X) \subseteq \underline{apr}_{E_1}(X) \subseteq X \subseteq \overline{apr}_{E_1}(X) \subseteq \overline{apr}_{E_2}(X).$$

That is, a finer equivalence relation induces a tighter pair of approximations. In terms of the accuracy measure, this implies:

$$E_1 \subseteq E_2 \implies \alpha_{E_2}(X) \leq \alpha_{E_1}(X).$$

The reverse is not necessarily true. From the values of the accuracy measure, one cannot tell if a pair of approximations is tighter than another pair, although one pair may have a smaller value.

In general, we may consider a nested sequence of m equivalence relations:

$$E_1 \subseteq E_2 \subseteq \dots \subseteq E_m.$$

The corresponding sequence of equivalence granules satisfies the condition:

$$[x]_{E_1} \subseteq [x]_{E_2} \subseteq \dots \subseteq [x]_{E_m}.$$

The nested sequence of equivalence relations produces a multi-level partitions of the universe. This leads to a simple multi-level granulation structure of the universe. Different granulations of the universe form a linear order. A partition is either a refinement or a coarsening of the other, although some granules in different levels may also be the same. The sequence of rough set approximations satisfies:

$$\begin{aligned} E_1 \subseteq E_2 \subseteq \dots \subseteq E_m \implies \\ \underline{apr}_{E_m}(X) \subseteq \dots \subseteq \underline{apr}_{E_2}(X) \subseteq \underline{apr}_{E_1}(X) \subseteq X \subseteq \\ \overline{apr}_{E_1}(X) \subseteq \overline{apr}_{E_2}(X) \subseteq \dots \subseteq \overline{apr}_{E_m}(X). \end{aligned}$$

It implies:

$$E_1 \subseteq E_2 \subseteq \dots \subseteq E_m \implies \alpha_{E_m}(X) \leq \dots \leq \alpha_{E_2}(X) \leq \alpha_{E_1}(X).$$

We thus obtain a nested sequence of rough set approximations, which may be viewed as a special type of stratified rough set approximations. As equivalence relations approaches to the identity relation $I = \{(x, x) \mid x \in U\}$, both lower and upper approximations approach to X , and the accuracy of approximation approaches to 1.

Example 3 A nested sequence of equivalence relations can be defined based on information provided by an information table. Let $A_1 \supseteq A_2 \supseteq \dots \supseteq A_m$ denote a nested sequence of sets of attributes. Suppose $E_i, 1 \leq i \leq m$, is the equivalence relation defined by the set of attribute A_i . We have a nested sequence of nested equivalence relations $E_1 \subseteq E_2 \subseteq \dots \subseteq E_m$. For Table 1, consider the sequence of subsets of attributions $\{A_1, A_2, A_3, A_4\}, \{A_1, A_2\}, \{A_1\}, \emptyset$. We have $I = E_{A_1 A_2 A_3 A_4} \subset E_{A_1 A_2} \subset E_{A_1} \subset E_\emptyset = U \times U$. The following multi-level granulation structure is obtained:

$$\begin{aligned} 4 &: \{\{a, b, c, d, e, f\}\}, \\ 3 &: \{\{a, b, c, d\}, \{e, f\}\}, \\ 2 &: \{\{a\}, \{b, c\}, \{d\}, \{e, f\}\}, \\ 1 &: \{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}, \{f\}\}. \end{aligned}$$

The top partition corresponds to the equivalence relation E_\emptyset . The nested rough set approximations of $+ = \{c, d, e, f\}$ in different level of granulations are given by:

level	lower apr	upper apr	accuracy
4	\emptyset	U	0
3	$\{e, f\}$	U	1/3
2	$\{d, e, f\}$	$\{b, c, d, e, f\}$	3/5
1	$\{c, d, e, f\}$	$\{c, d, e, f\}$	1

In a higher level with coarser granulation, one obtains less accurate rough set approximations. One may search the layered granulations to find the suitable granulation for approximating $+$. \square

The multi-level granulation structures defined by a nested sequence of partitions can be easily generalized. One may consider a nested sequence of tolerance relations [22, 35]. This leads to a multi-level granulation structure, in which each level of granulation is a covering of the universe. In an information table, one may consider all possible subsets of attributes. The resulting multi-level granulation structure is characterized by a lattice of partitions [5]. A nested sequence partitions is a special type of lattice.

In the discussion so far, we have used information tables for the representation of objects. The similarities between objects used for granulations are derived from their attribute values. This provides us with a more concrete interpretation of granulation and approximation. In the following subsections, we consider other approaches for the representation of similarities between objects.

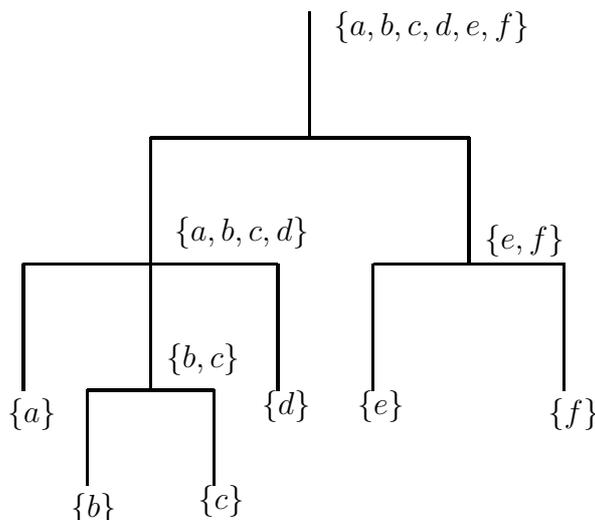


Figure 1: A Hierarchical Granulation

3.2 Stratified rough set approximations induced by hierarchies

Similarities between objects can be conveniently represented by using a hierarchy on a universe. It can be described by a tree structure such that each node represents a cluster or granule. Figure 1 is an example of a hierarchy. For simplicity, we assume that the root is the entire universe, and the leaves consist of only singleton subsets. We further assume that granules containing x are distinct at different levels. Conceptually, a hierarchy may be viewed as a successive top-down decomposition of a universe U . The root is divided into a family of pairwise disjoint clusters. That is, the children clusters of the root form a partition of the root. Each cluster is further divided into smaller disjoint clusters. Alternatively, a hierarchy may also be viewed as a successive bottom-up combination of smaller clusters to form larger clusters. In a hierarchy, all elements of a cluster at a lower level are included in every node between that cluster and the root, which form a sequence of nested clusters.

From a hierarchy, one may obtain two partition based multi-level granulation structures of the universe. One leads to a nested sequence of partitions, and the other leads to a lattice of partitions.

Given a m -level hierarchy with the root at level m , one can derive a nested sequence of equivalence relations such that $E_1 = I$ and $E_m = U \times U$. For $x \in U$,

suppose $\{x\} = F_{k(x)}(x) \subset \dots \subset F_m(x) = U, 1 \leq k(x) \leq m$, is the nested sequence of clusters containing x . At level $1 \leq l \leq m$, the equivalence granule containing x is given by $F_l(x)$ if $l \geq k(x)$, otherwise, it is given by $F_{k(x)}(x)$. Obviously, at the highest level, we obtain the relation $E_m = U \times U$, and at the lowest level, we obtain the identity relation I . A main disadvantage of such a characterization of hierarchy is that one considers a nested sequence of rough set approximations.

Consider now the set of all granules in a hierarchy:

$$G = \{X \subseteq U \mid X \text{ is a node in the hierarchy}\}. \quad (14)$$

For any two granules $X, Y \in G$, we have $X \cap Y = X$, $X \cap Y = Y$, or $X \cap Y = \emptyset$. We can select a subset of G to form a partition of the universe. The set of all partitions constructed from elements of G is denoted by $P(G)$, and the corresponding set of equivalence relations is denoted by $E(G)$. By assumption, $I, U \times U \in E(G)$. The family of equivalence relations $E(G)$ is closed under set intersection and union. For any two equivalence relations $E_1, E_2 \in E(G)$, we have $E_1 \cap E_2, E_1 \cup E_2 \in E(G)$. In general, the union of two arbitrary equivalence relations is not necessarily an equivalence relation. For our case, the special properties of elements of G guarantee that $E(G)$ is closed under set union. The set $E(G)$ is a bounded lattice whose order relation is the standard set inclusion, and whose meet and join are set intersection and union.

If the granulation structure induced by $E(G)$ is used, we obtain stratified rough set approximations. They carry over the structure of $E(G)$. In other words, the stratified rough set approximations form a lattice. For any pair of equivalence relations in $E(G)$, the following properties hold:

$$\begin{aligned} \underline{apr}_{E_1 \cup E_2}(X) &\subseteq \frac{\underline{apr}_{E_1}(X)}{\underline{apr}_{E_2}(X)} \subseteq \underline{apr}_{E_1 \cap E_2}(X) \subseteq X \subseteq \\ \overline{apr}_{E_1 \cap E_2}(X) &\subseteq \frac{\overline{apr}_{E_1}(X)}{\overline{apr}_{E_2}(X)} \subseteq \overline{apr}_{E_1 \cup E_2}(X). \end{aligned}$$

One can apply the same argument to more than two equivalence relations from $E(G)$. The induced approximations also produce a lattice with the standard set inclusion as its order relation.

Example 4 Consider the 4-level hierarchy for the universe $U = \{a, b, c, d, e, f\}$ given by Figure 1. For element a , the family of granules is given by:

$$\{a\} = F_2(a) \subset \{a, b, c, d\} = F_3(a) \subset U = F_4(a).$$

Similarly, we can find the nested sequences for all other elements. Thus, we have a family of 4-level partitions:

$$\begin{aligned}
 4 &: \{\{a, b, c, d, e, f\}\}, \\
 3 &: \{\{a, b, c, d\}, \{e, f\}\}, \\
 2 &: \{\{a\}, \{b, c\}, \{d\}, \{e\}, \{f\}\}, \\
 1 &: \{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}, \{f\}\}.
 \end{aligned}$$

From the layered granulations of the universe, one can obtain a nested sequence of rough set approximations.

From the hierarchy, the set of granules is given by:

$$G = \{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}, \{f\}, \{b, c\}, \{e, f\}, \{a, b, c, d\}, U\}.$$

From G , we can construct the set of all possible partitions:

$$\begin{aligned}
 \pi_1 &: \{U\}, \\
 \pi_2 &: \{\{a, b, c, d\}, \{e, f\}\}, \\
 \pi_3 &: \{\{a, b, c, d\}, \{e\}, \{f\}\}, \\
 \pi_4 &: \{\{a\}, \{b, c\}, \{d\}, \{e, f\}\}, \\
 \pi_5 &: \{\{a\}, \{b, c\}, \{d\}, \{e\}, \{f\}\}, \\
 \pi_6 &: \{\{a\}, \{b\}, \{c\}, \{d\}, \{e, f\}\}, \\
 \pi_7 &: \{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}, \{f\}\}.
 \end{aligned}$$

Figure 2 shows the relationships between these partitions, the structure of granulations using the family of equivalence relations, and the stratified rough set approximations of set $\{a, b, e\}$. \square

In a hierarchy, one typically associates a name with a cluster such that elements of the cluster are instances of the named category or concept [3, 8]. Suppose U is the domain of an attribute in a database. A hierarchical clustering of attribute values produces a concept hierarchy [2]. A name given to a cluster in a higher level is more general than a name given to a cluster in a lower level, while the latter is more specific than the former. The notion of concept hierarchy has been used in data mining for discovering various levels of association rules [2]. Partitions in higher levels of the partition lattice may be viewed as generalization of partitions in lower levels, while partitions in lower levels as specialization of partitions in higher levels. A similar structure was discussed by Hadjimichael and Wasilewska [1] for the study of a hierarchical model for information generalization. Conceptually, some data mining methods may be viewed as a searching process in the lattice of partitions induced by a hierarchy [1, 2].

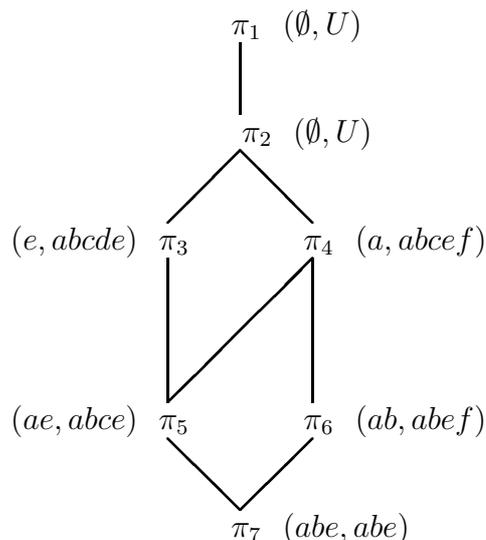


Figure 2: A lattice granulation structure and the corresponding stratified rough set approximations (the set $\{a, b, e\}$ is simply written as abe .)

3.3 Stratified rough set approximations induced by neighborhood systems

The concept of neighborhood systems was originally introduced by Sierpiński and Krieger [23] for the study of F chet (V)spaces. Lin [6, 7] adopted the notion for describing similarities between objects in database systems. Yao [30] used the notion for granular computing by focusing on the granulation structures induced by neighborhood systems.

For an element x of a finite universe U , one associates with it a subset $n(x) \subseteq U$ called the neighborhood of x . Intuitively speaking, elements in a neighborhood of an element are somewhat indiscernible or at least not noticeably distinguishable from x . A neighborhood of x may or may not contain x . A neighborhood of x containing x is called a reflexive neighborhood. We are only interested in reflexive neighborhoods of x to accommodate the intuitive interpretation of neighborhoods. A neighborhood system $NS(x)$ of x is a nonempty family of neighborhoods of x . Distinct neighborhoods of x consist of elements having different types of, or various degrees of, similarity to x . A neighborhood system is reflexive, if every neighborhood in it is reflexive. Let $NS(U)$ denote the collection of neighborhood systems for all elements in U . It determines a F chet (V)space, written $(U, NS(U))$. There is no additional

requirements on neighborhood systems.

Neighborhood systems can be used to describe more general types of similarities between elements of a universe [7, 30]. All previously used similarities can be explained in terms of neighborhood systems. A binary relation can be interpreted in terms of 1-neighborhood systems, in which each neighborhood system contains only one neighborhood [29]. More precisely, the neighborhood system of x is given by $NS(x) = \{(x)_R\}$. If R is a reflexive relation, one obtains a reflexive neighborhood system which is the covering U/R . If R is an equivalence relation, the successor neighborhood $(x)_R$ is the equivalence class containing x , and the neighborhood system is the partition U/R . A family of nested sequence of m binary relations defined a nested neighborhood system $\{(x)_{R_i} \mid 1 \leq i \leq m\}$. In a hierarchy, all clusters containing x may be used to form a nested neighborhood system of x .

In a neighborhood system, different neighborhoods represent different types or degrees of similarity. Such information should be taken into consideration. By extending the method for building a partition lattice from a hierarchy, we can constructed a family of coverings from a neighborhood system of the universe. Instead of using *all* neighborhoods, each covering is obtained by selecting *one* particular neighborhood for each element, i.e.,

$$C = (n(x), n(y), \dots, n(z)), \quad (15)$$

where $n(x) \in NS(x), n(y) \in NS(y), \dots, n(z) \in NS(z)$ for $x, \dots, y, z \in U$. In this way, we transform a neighborhood system into a family of 1-neighborhood systems $FC(U)$. An order relation \preceq on $FC(U)$ can be defined as follows, for $C_1, C_2 \in FC(U)$,

$$C_1 \preceq C_2 \iff n_{C_1}(x) \subseteq n_{C_2}(x), \quad \text{for all } x \in U. \quad (16)$$

The covering C_1 is finer than C_2 , or C_2 is coarser than C_1 . For each granule in C_2 produced by x , the granule in C_1 produced by x is at least as small as the former. It can be verified that \preceq is reflexive, transitive, and anti-symmetric. In other words, \preceq is a partial order, and the set $FC(U)$ is a poset. It is not necessarily a lattice. Thus, we have obtained a family of multi-level coverings, which in turn produces multi-level granulations of the universe.

For each covering $C \in FC(U)$, we can define three pairs of lower and upper approximations by using equations (9), (10), and (11). With respect to the poset $FC(U)$, we obtain multi-level approximations. For reflexive neighborhood system,

approximations in various levels satisfy the conditions:

$$\begin{aligned} C_1 \preceq C_2 &\implies \underline{apr}''_{C_2}(X) \subseteq \underline{apr}''_{C_1}(X) \subseteq X \subseteq \overline{apr}''_{C_1}(X) \subseteq \overline{apr}''_{C_2}(X), \\ C_1 \preceq C_2 &\implies \underline{apr}_{C_2}(X) \subseteq \underline{apr}_{C_1}(X) \subseteq X \subseteq \overline{apr}_{C_1}(X) \subseteq \overline{apr}_{C_2}(X). \end{aligned} \quad (17)$$

A finer covering C_1 produces a better approximation than a coarser covering C_2 . In other words, both approximations $(\underline{apr}'', \overline{apr}'')$ and $(\underline{apr}, \overline{apr})$ define approximation structures characterized by posets. On the other hand, the approximation $(\underline{apr}', \overline{apr}')$ does not induce such a structure.

Example 5 Consider the neighborhood systems on a universe $U = \{a, b, c, d\}$:

$$\begin{aligned} \text{NS}(a) &= \{\{a\}, \{a, b\}\}, \\ \text{NS}(b) &= \{\{a, b\}, \{a, b, c\}\}, \\ \text{NS}(c) &= \{\{c\}\}, \\ \text{NS}(d) &= \{\{c, d\}, \{b, d\}\}. \end{aligned}$$

From these neighborhood systems, we obtain eight coverings:

$$\begin{aligned} C_1 : & (\{a\}, \{a, b\}, \{c\}, \{c, d\}), \\ C_2 : & (\{a, b\}, \{a, b\}, \{c\}, \{c, d\}), \\ C_3 : & (\{a\}, \{a, b, c\}, \{c\}, \{c, d\}), \\ C_4 : & (\{a, b\}, \{a, b, c\}, \{c\}, \{c, d\}), \\ C_5 : & (\{a\}, \{a, b\}, \{c\}, \{b, d\}), \\ C_6 : & (\{a, b\}, \{a, b\}, \{c\}, \{b, d\}), \\ C_7 : & (\{a\}, \{a, b, c\}, \{c\}, \{b, d\}), \\ C_8 : & (\{a, b\}, \{a, b, c\}, \{c\}, \{b, d\}). \end{aligned}$$

Figure 3 shows the granulation structure and the stratified rough set approximation for the subset $\{a, b, d\}$. □

4 Conclusion

In this paper, we investigate some fundamental issues of granulation and approximation in the context of rough set theory. Our discussion is based on the notion of *similarity* that represents relationships between elements of a universe. Depending

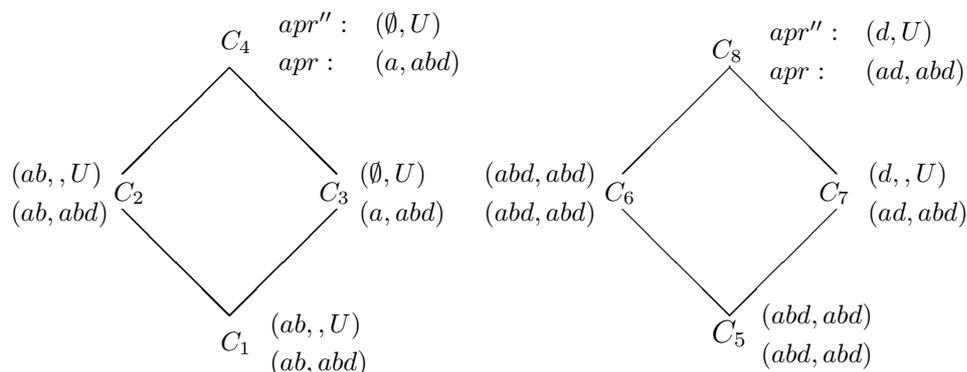


Figure 3: A poset granulation structure and the corresponding stratified rough set approximations (the set $\{a, b, d\}$ is simply written as abd .)

on the various interpretations of similarity, different granulation structures are examined. We start from two simple granulation structures. One structure is defined by an equivalence relation, which lead to a partition of the universe. In this case, the standard rough set approximation is used. The other structure is defined by a reflexive binary relation that induces a covering of the universe. Three generalized rough set approximations are proposed. From the two simple one-level granulation structures, we can study more general granulation structures characterized by many levels of simple granulation structures. In particular, we analyze multi-level granulation structures induced by hierarchies of the universe and neighborhood systems. The former leads to partition based granulation structures, and the latter leads to covering based granulation structures. With the multi-level structure, we examine stratified rough set approximations.

Granulation structures and the corresponding approximation structures introduced in this paper provide a starting point for further study of granulation and approximation. Investigations in this direction may produce interesting and useful results.

References

- [1] Hadjimichael, M. and Wasilewska, A. A hierarchical model for information generalization, *Proceedings of 1998 Joint Conference on Information Sciences*, Vol. III, 306-309, 1998.

- [2] Han, J.W., Cai, Y. and Cercone, N., 1993, Data-driven discovery of quantitative rules in relational databases, *IEEE Transactions on Knowledge and Data Engineering*, **5**, 29-40.
- [3] Jardine, N. and Sibson, R., 1971, *Mathematical Taxonomy*, Wiley, New York.
- [4] Klir, G.J. Basic issues of computing with granular probabilities, *Proceedings of 1998 IEEE International Conference on Fuzzy Systems*, 101-105, 1998.
- [5] Lee, T.T. An information-theoretic analysis of relational databases – part I: data dependencies and information metric, *IEEE Transactions on Software Engineering*, **SE-13**, 1049-1061, 1987.
- [6] Lin, T.Y. Neighborhood systems and approximation in relational databases and knowledge bases, *Proceedings of the 4th International Symposium on Methodologies of Intelligent Systems*, 1988.
- [7] Lin, T.Y. Granular computing on binary relations I: data mining and neighborhood systems, II: rough set representations and belief functions, in: *Rough Sets in Knowledge Discovery 1*, Polkowski, L. and Skowron, A. (Eds.), Physica-Verlag, Heidelberg, 107-140, 1998.
- [8] Lin, T.Y. and Hadjimichael, M., 1996, Non-classificatory generation in data mining, *Proceedings of the 4th International Workshop on Rough Sets, Fuzzy Sets, and Machine Discovery*, 404-411.
- [9] Lipski, W. Jr. On databases with incomplete information, *Journal of the ACM*, **28**, 41-70, 1981.
- [10] Marczewski, E. and Steinhaus, H. On a certain distance of sets and the corresponding distance of functions, *Colloquium Mathematicum*, **6**, 319-327, 1958.
- [11] Marek, W. and Rasiowa, H. Gradual approximating sets by means of equivalence relations, *Bulletin of Polish Academy of Sciences, Mathematics*, **35**, 233-238, 1987.
- [12] Orłowska, E. Logic of nondeterministic information, *Studia Logica*, **XLIV**, 93-102, 1985.
- [13] Pawlak, Z. Information systems: theoretical foundations, *Information Systems*, **6**, pp. 205-218, 1981.

- [14] Pawlak, Z. Rough sets, *International Journal of Computer and Information Sciences*, **11**, 341-356, 1982.
- [15] Pawlak, Z. *Rough Sets, Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, 1991.
- [16] Pawlak, Z. Granularity of knowledge, indiscernibility and rough sets, *Proceedings of 1998 IEEE International Conference on Fuzzy Systems*, 106-110, 1998.
- [17] Polkowski, L. On convergence of rough sets, in: *Intelligent Decision Support: Handbook of Applications and Advances of Rough Set Theory*, Kluwer, Dordrecht, 305-311, 1992.
- [18] Polkowski, L. Mathematical morphology of rough sets, *Bulletin of Polish Academy of Sciences, Mathematics*, **41**, 241-273, 1993.
- [19] Polkowski, L. Rough set approach to mathematical morphology: approximate compression of data, *Proceedings of Seventh International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems*, 1183-1189, 1998.
- [20] Polkowski, L. Approximate mathematical morphology: rough set approach, in: *Fuzzy Sets and Rough Sets in Soft Computing*, Springer Singapore, 151-162, 1999.
- [21] Polkowski, L. and Skowron, A. Towards adaptive calculus of granules, *Proceedings of 1998 IEEE International Conference on Fuzzy Systems*, 111-116, 1998.
- [22] Pomykala, J.A. A remark on the paper by H. Rasiowa and W. Marek: "Gradual approximating sets by means of equivalence relations", *Bulletin of Polish Academy of Sciences, Mathematics*, **36**, 509-512, 1988.
- [23] Sierpiński, W. and Krieger, C. *General Topology*, University of Toronto, Toronto, 1956.
- [24] Skowron, A. and Stepaniuk, J. Information granules and approximation spaces, manuscript, 1998.
- [25] Slowinski, R. and Vanderpooten, D. Similarity relation as a basis for rough approximations, in: *Advances in Machine Intelligence & Soft-Computing*, Wang, P.P. (Ed.), Department of Electrical Engineering, Duke University, Durham, North Carolina, 17-33, 1997.

- [26] Vakarelov, D. A modal logic for similarity relations in Pawlak knowledge representation systems, *Fundamenta Informaticae*, **15**, 61-79, 1991.
- [27] Yager, R.R. and Filev, D. Operations for granular computing: mixing words with numbers, *Proceedings of 1998 IEEE International Conference on Fuzzy Systems*, 123-128, 1998.
- [28] Y.Y. Yao, Two views of the theory of rough sets in finite universes, *International Journal of Approximate Reasoning*, **15**, 291-317, 1996.
- [29] Yao, Y.Y. Relational interpretations of neighborhood operators and rough set approximation operators, *Information Sciences*, **111**, 239-259, 1998.
- [30] Yao, Y.Y. Granular computing using neighborhood systems, in: *Advances in Soft Computing: Engineering Design and Manufacturing*, Roy, R., Furuhashi, T., and Chawdhry, P.K. (Eds.), Springer-Verlag, London, 539-553, 1999.
- [31] Yao, Y.Y. Stratified rough sets and granular computing, *Proceedings of the 18th International Conference of the North American Fuzzy Information Processing Society*, IEEE Press, pp. 800-804, 1999.
- [32] Yao, Y.Y. Rough sets, neighborhood systems, and granular computing, *Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering, Edmonton*, IEEE Press, pp. 1553-1558, 1999.
- [33] Yao, Y.Y. Granular computing: basic issues and possible solutions, *Proceedings of the 5th Joint Conference on Information Sciences*, 186-189, 1999.
- [34] Yao, Y.Y. and Lin, T.Y. Generalization of rough sets using modal logic, *Intelligent Automation and Soft Computing, an International Journal*, **2**, 103-120, 1996.
- [35] Yao, Y.Y. and Lin, T.Y. Graded rough set approximations based on nested neighborhood systems, *Proceedings of 5th European Congress on Intelligent Techniques & Soft Computing*, 196-200, 1997.
- [36] Yao, Y.Y. and Noroozi, N. A unified model for set-based computations, in: *Soft Computing*, Lin, T.Y. and Wildberger, A.M. (Eds.), The Society for Computer Simulation, San Diego, pp. 252-255, 1995.

- [37] Yao, Y.Y., Wong, S.K.M., and Lin, T.Y. A review of rough set models, in: Lin, T.Y. and Cercone, N. (Eds.), *Rough Sets and Data Mining: Analysis for Imprecise Data*, Academic Publishers, Boston, 47-75, 1997.
- [38] Zadeh, L.A. Fuzzy sets and information granularity, in: *Advances in Fuzzy Set Theory and Applications*, Gupta, N., Ragade, R. and Yager, R. (Eds.), North-Holland, Amsterdam, 3-18, 1979.
- [39] Zadeh, L.A. Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets and Systems*, **19**, 111-127, 1997.
- [40] Zadeh, L.A. and Kacprzyk, J. (Eds.), *Computing with Words in Information/Intelligent Systems*, Volume 1 and 2, Physica-Verlag, Heidelberg, 1999.