Study on Rough Sets and Fuzzy Sets in Constructing Intelligent Information System

Rolly INTAN

PETRA PRESS

Study on Rough Sets and Fuzzy Sets in Constructing Intelligent Information System
By: Rolly INTAN

Published by:
Institute for Research and Community Outreach
Petra Press
Petra Christian University
Jl. Siwalankerto 121-131, Surabaya 60236
INDONESIA
http://lppm.petra.ac.id

ISBN 978-602-71225-8-1

Preface

This book is an extending part of my doctoral dessertation providing a research in constructing intelligent information system dealing with rough sets, fuzzy sets and granular computing. Uncertainty was involved and connected to every aspect of human life. The most fundamental aspect of this connection is obviously shown in human communication. Naturally, human communication is built on perception based information. Perceptions are intrinsic aspect in uncertainty based information. In this case, information may be incomplete, imprecise, fragmentary, not fully reliable, vague, contradictory, or deficient in some other way. Generally, these various information deficiencies may express different types of uncertainty. Many theories were proposed to express and process the types of uncertainty such as probability, possibility, fuzzy sets, rough sets, chaos theory and so on.

It is necessary to construct a computer-based information system called *intelligent information system* that can process uncertainty-based information. In the future, computers are expected to be able to make communication with human in the level of perception. This book extends and generalizes existing theory of rough set, fuzzy sets and granular computing for the purpose of constructing the intelligent information system.

I realized what written in this book only shared a small part of research topics in rough sets, fuzzy sets and granular computing in the relation to construct perception based information. However, I hope that this book will be a valuable reference especially for undergraduate as well as graduate students who are interested to study and do a research in the topics.

Finally, I would like to express my gratitude to my doctoral advisor, Prof. Dr. Masao Mukaidono, for his support and guidance throughout the research during my doctoral study at Meiji University, Tokyo, Japan from 2000 to 2003. I would also like to extend my gratitude to my family. Without their encouragement and sacrifice, I would not complete my doctoral study on time successfully.

Surabaya, October 2015
Rolly INTAN

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since human being is not an omniscient and omnipotent being, we are actually living in an uncertain world. Uncertainty was involved and connected to every aspect of human life as a quotation from Albert Einstein said:

> "*As far as the laws of mathematics refer to reality, they are not certain. And as far as they are certain, they do not refer to reality.*"[1]

The most fundamental aspect of this connection is obviously shown in human communication. Naturally, human communication is built on *the perception*[1]*-based information* instead of *measurement-based information* in which perceptions play a central role in human cognition [Zadeh, 2000]. For example, it is naturally said in our communication that "My house is far from here." rather than let say "My house is 12,355 m from here". Perception-based information is a generalization of measurement-based information, where perception-based information such as "John is excellent." is hard to represent by measurement-based version. Perceptions express human subjective view. Consequently, they tend to lead up to misunderstanding. Measurements then are needed such as defining units of length, time, etc., to provide objectivity as a means to overcome misunderstanding. Many measurers were invented along with

[1] In psychology, *perception* is understood as a process of translating sensory stimulation into an organized experience.

their methods and theories of measurement. Hence, human cannot communicate with measurers including computer as a product of measurement era unless he uses measurement-based information.

Perceptions are intrinsic aspect in uncertainty-based information. In this case, information may be incomplete, imprecise, fragmentary, not fully reliable, vague, contradictory, or deficient in some other way.

Generally, these various information deficiencies may express different types of uncertainty. It is necessary to construct a computer-based information system called *intelligent information system* that can process uncertainty-based information. In the future, computers are expected to be able to make communication with human in the level of perception.

Many theories were proposed to express and process the types of uncertainty such as probability, possibility, fuzzy sets, rough sets, chaos theory and so on. This book extends and generalizes existing theory of rough set, fuzzy sets and granular computing for the purpose of constructing intelligent information system. The structure of this book is the following: In Chapter 2, types of uncertainty in the relation to fuzziness, probability and evidence theory (belief and plausibility measures) are briefly discussed. Rough set regarded as another generalization of crisp set is considered to represent rough event in the connection to the probability theory. Special attention will be given to formulation of *fuzzy conditional probability relation* generated by property of conditional probability of fuzzy event. Fuzzy conditional probability relation then is used to represent similarity degree of two fuzzy labels.

Generalization of rough set induced by fuzzy conditional probability relation in terms of covering of the universe is given in Chapter 3. In the relation to fuzzy conditional probability relation, it is necessary to consider an interesting mathematical relation called *weak fuzzy similarity relation* as a generalization of *fuzzy similarity relation* proposed by Zadeh [1995]. Fuzzy rough set and generalized fuzzy rough set are proposed along with the generalization of rough membership function. Their properties are examined. Some applications of these methods in information system such as α-redundancy of object and dependency of domain attributes are discussed.

In addition, multi rough sets based on multi-context of attributes in the presence of multi-contexts information system is defined and proposed in Chapter 4. In the real application, depending on the context, a given object may have different values of attributes. In other words, set of attributes might be represented based on different context, where they may provide different values for a given object. Context can be viewed as background or situation in which somehow it is necessary to group some attributes as a subset of attributes and consider the subset as a context.

Finally, Chapter 5 summarizes all discussed in this book and puts forward some future topics of research.

Chapter 2

Probability, Fuzziness, Rough and Evidence Theory

2.1 Introduction

Since the appearance of the first article on fuzzy sets proposed by Zadeh in 1965, the relationship between probability and fuzziness in representing uncertainty has been an object of debate among many people. The main problem is whether or not probability theory by itself is sufficient for dealing with uncertainty. This question has been discussed at length in many papers such as written by Nguyen [Nguyen, 1977], Kosko [Kosko, 1990], Zadeh [Zadeh, 1968, 1995] and so on.

In this chapter, again the process of perception performed by human being is used to simply understand the relationship between probability and fuzziness. In the process of perception, subject (human, computer, robot, etc.) tries to recognize and describe a given object (anything such as human, plant, animal, event, condition, etc.). To perform perception successfully, subject needs adequate knowledge. On the other hand, object needs a clear definition. However, human (as subject) does not know what happen in the future and also has limited knowledge. In other words, human is not omniscient being. In this case, subject is in a non-deterministic situation in performing a perception. On the other hand, mostly objects (shape, feel, mentality, etc.) cannot usually be defined clearly. Therefore, the process of perception turns into uncertainty.

To summarize the relation between subject and object in the process of perception, there are four possible situations as follows.

(a) If subject has sufficient knowledge and object has clear definition, it comes to be a *certainty*.

(b) If subject has sufficient knowledge and object has unclear definition, it comes to be *fuzziness*. In general, fuzziness, called deterministic uncertainty, may happen in the situation when one is subjectively able to determine or describe a given object, although somehow the object does not have a certain or clear definition. For example, a man describes a woman as a *pretty* woman. Obviously definition of a pretty woman is unclear, uncertain and subjective. The man however is convinced of what he describes as a pretty woman.

(c) If subject does not have sufficient knowledge and object has clear definition, it comes to be *randomness*. Randomness is usually called non-deterministic uncertainty because subject cannot determine or describe a given object even though the object has clear definition. Here, probability exists for measuring a random experiment. For example, in throwing a dice, even though there are six definable and certain possibilities of outcome, one however cannot assure the outcome of dice. Still another example, because of his limited knowledge, for instance, one cannot assure to choose a certain answer in a multiple choice problem in which there are 4 possible answers, but only one answer is correct.

(d) If subject does not have sufficient knowledge and object has unclear definition, it comes to be a *probability of fuzzy event* [Zadeh, 1968]. In this situation, both probability and fuzziness are combined. For example, how to predict the ill-defined event: *"Tomorrow will be a warm day"*. Talking about tomorrow means talking about the future in which subject cannot determine what happen in the future. The situation should be dealt by probability. However, *warm* is an ill-defined event (called fuzzy event). Therefore, it comes to be a *probability of fuzzy event.*

From these four situations, it is obviously seen that probability and fuzziness work in different areas of uncertainty and that probability theory by itself is not sufficient for especially dealing with ill-defined event. Instead, probability and fuzziness must be regarded as a complementary tool.

In probability, set theory is used to provide a language for modeling and describing random experiments. In (classical) set theory, subsets of the sample space of an experiment are referred to as *crisp events*. Fuzzy set theory, proposed by Zadeh in 1965, is considered as a generalization of (classical) set theory in which fuzzy set is to represent deterministic uncertainty by a class or classes which do not possess sharply defined boundaries [Zadeh, 1990]

By fuzzy set, an ill-defined event, called fuzzy event, can be described in the presence of probability theory providing *probability of fuzzy event* [Zadeh, 1968] in which fuzzy event might be regarded as a generalization of crisp event. Conditional probability as an important property in probability theory for inference rule can be extended to conditional probability of fuzzy event. In the situation of uniform probability distribution, conditional probability of fuzzy event can be simplified to be what we call *fuzzy conditional probability relation* as proposed in [Intan, Mukaidono, 2000a, 2000c] for dealing with similarity of two fuzzy labels (sets).

Similarly, rough set theory generalizes classical set theory by studying sets with imprecise boundaries. A rough set [Pawlak, 1982], characterized by a pair of lower and upper approximations, may be viewed as an approximate representation of a given crisp set in terms of two subsets derived from a partition on the universe [Klir, Yuan, 1995], [Komorowski, Pawlak, Polkowski, Skowron, 1999], [Pawlak 1982], [Yao, 1996]. By rough set theory, a rough event is proposed to represent two approximate events, namely lower and upper approximate events, in the presence of probability theory providing *probability of rough event*. Therefore, rough event might be considered as approximation of a given crisp event. Moreover, probability of rough event gives semantic formulation of interval probability. Formulation of interval probability is useful in order to represent the worst and the best cases in decision making process. In this chapter, special attention will be given to conditional probability of rough event providing several combinations of formulation and properties.

In addition, a generalized fuzzy rough set as proposed in [Intan, Mukaidono, 2002f], [Intan, Mukaidono, 2002k] (see Chapter 3) is an approximation of a given fuzzy set on a given fuzzy covering. Since fuzzy set generalizes crisp set and fuzzy covering generalizes crisp partition, the generalized fuzzy rough set is considered as the most generalization of fuzzy set and rough set as well as rough fuzzy set and fuzzy rough set as proposed in [Dubois, Prade, 1990]. Thus, by the generalized fuzzy rough set, a generalized fuzzy-rough event is proposed providing probability of the generalized fuzzy-rough event. The generalized fuzzy-rough event is represented in four approximate fuzzy events, namely lower minimum, lower maximum, upper minimum and upper maximum fuzzy events.

Finally, this chapter shows and discusses relation among belief-plausibility measures (evidence theory), lower-upper approximate probability (probability of rough events), classical probability measures, probability of fuzzy events and probability of generalized fuzzy-rough events.

2.2 Probability of Fuzzy Event

Probability theory is based on the paradigm of a random experiment; that is, an experiment whose outcome cannot be predicted with certainty, before the experiment is run. In other words, as discussed in the previous section, probability is based on that a subject has no sufficient knowledge in certainly predicting (determining) outcome of an experiment. In probability, set theory is used to provide a language for modeling and describing random experiments. The sample space of a random experiment corresponds to universal set. In (classical) set theory, subsets of the sample space of an experiment are referred to be crisp events.

In order to represent an ill-defined event, crisp event must be generalized to fuzzy event in which fuzzy set is used to represent fuzzy event. Formally, probability of fuzzy event is defined as the following [Zadeh, 1968]:

Definition 2.2.1 Let (U,F,P) be a probability space in which U is the sample space, F is sigma algebra of events and P is a probability measure over U. Then, a fuzzy event $A \in F$ is a fuzzy set A on U whose membership function, $\mu_A: U \rightarrow [0, 1]$. The probability of fuzzy event A is defined by:

- *continuous sample space:*

$$P(A) = \int_U \mu_A(u)dP = \int_U \mu_A(u) \cdot p(u)du \quad (2.1)$$

- *discrete sample space:*

$$P(A) = \sum_U \mu_A(u) \cdot p(u) \quad (2.2)$$

where $p(u)$ is probability distribution function of element $u \in U$.

For example, it is given a sentence "John ate a few eggs for breakfast", and we do not know exactly how many eggs John ate for breakfast. In this case, arbitrarily given probability distribution function of "John ate $u \in U$ egg(s) for breakfast" is shown in Table 2.1.

Table 2.1. Probability Distribution of u

u	1	2	3	4	5	6	• • •
$p(u)$	0.33	0.27	0.2	0.13	0.07	0	• • •

"*a few*" is a fuzzy label that also means a fuzzy event as arbitrarily given by the following fuzzy set: $\mu_{afew} = \{1/1, 0.6/2, 0.2/3\}$, where $\mu_{afew}(2) = 0.6$. By Definition 2.2.1, probability of "John ate a few eggs for breakfast", denoted by P (*a few*), is calculated as:

$$P(a\ few) = 1 \times 0.33 + 0.6 \times 0.27 + 0.2 \times 0.2 = 0.532.$$

There are several basic concepts relating to fuzzy sets. For A and B are two fuzzy sets on U [Zadeh, 1990],

Equality: $A = B \Leftrightarrow \mu_A(u) = \mu_B(u), \forall u,$

Containment: $A \subset B \Leftrightarrow \mu_A(u) \leq \mu_B(u), \forall u,$

Complement: $B = \neg A \Leftrightarrow \mu_B(u) = 1 - \mu_A(u), \forall u,$

Union: $\mu_{A \cup B}(u) = \max[\mu_A(u), \mu_B(u)],$

Intersection: $\mu_{A \cap B}(u) = \min[\mu_A(u), \mu_B(u)],$

Product: $\mu_{AB}(u) = \mu_A(u) \cdot \mu_B(u),$

Sum: $\mu_{A \oplus B}(u) = \mu_A(u) + \mu_B(u) - \mu_A(u) \cdot \mu_B(u).$

Obviously, it can be proved that probability of fuzzy event satisfies some properties: for A and B are two fuzzy sets on U,

(1) $A \subset B \Rightarrow P(A) \leq P(B),$

(2) $P(A \cup B) = P(A) + P(B) - P(A \cap B),$

(3) $P(A \oplus B) = P(A) + P(B) - P(A \cdot B),$

(4) $P(A \cup \neg A) \leq 1,$

(5) $P(A \cap \neg A) \geq 0.$

(1), (2) and (3) show that probability of fuzzy event satisfies *monotonicity* and *additivity* axioms of union as well as sum operation, respectively. However, it does not satisfy law of excluded middle and law of non-contradiction as shown in (4) and (5).

We turn next to notion of conditional probability of fuzzy events. Conditional Probability of an event is probability of the event occurring given that another event has already occurred. The relationship between conditional and unconditional probability satisfies the following equation:

$$P(A|B) = P(A \cap B)/P(B),$$

where suppose B is an event such that $P(B) \neq 0$.

In discrete sample space, conditional probability of fuzzy event might be defined as follow: for A and B are two fuzzy sets on U,

$$P(A \mid B) = \frac{\sum_U \min[\mu_A(u), \mu_A(u)] \cdot p(u)}{\sum_U \mu_B(u) \cdot p(u)}, \forall u \in U, \qquad (2.3)$$

where $\sum_U \mu_B(u) \cdot p(u) > 0$. Some properties are satisfied in conditional probability of fuzzy event: for A and B are two fuzzy sets on U,

(1) Normalization: $P(A|B) + P(\neg A|B) \geq 1$,

(2) Total Probability; If $\{B_k \mid k \in \mathrm{N}_n\}$ are crisp, pairwise disjoint and exhaustive events,

i.e., $P(B_i \cap B_j) = 0$ for $i \neq j$ and $\cup B_k = U$, then:

$$P(A) = \sum_k P(B_k) \cdot P(A \mid B_k),$$

(3) Bayes Theorem:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}.$$

Also, the relationship between A and B in conditional probability of fuzzy event can be represented into three conditions:

(a) positive correlation:

$P(A|B) > P(A) \Leftrightarrow P(B|A) > P(B) \Leftrightarrow P(A \cap B) > P(A) \times P(B)$,

(b) negative correlation:

$P(A|B) < P(A) \Leftrightarrow P(B|A) < P(B) \Leftrightarrow P(A \cap B) < P(A) \times P(B)$,

(c) independent correlation:

$P(A|B) = P(A) \Leftrightarrow P(B|A) = P(B) \Leftrightarrow P(A \cap B) = P(A) \times P(B)$.

In uniform distribution, probability distribution function, $p(u) = 1/|U|$, is regarded as a constant variable. Therefore, conditional probability of fuzzy event A given B is defined more simply without $p(u)$ by:

$$P(A \mid B) = \frac{\sum_U \min[\mu_A(u), \mu_A(u)]}{\sum_U \mu_B(u)}, \forall u \in U, \quad (2.4)$$

In [Intan, Mukaidono, 2000a], [Intan, Mukaidono, 2000c], we used the formula to calculate degree of similarity relationship between two fuzzy labels (sets) and called it, *fuzzy conditional probability relation*.

2.3 Probability of Rough Event

Rough set is another generalization of crisp set by studying sets with imprecise boundary. A rough set, characterized by a pair of lower and upper approximations, may be viewed as an approximate representation of a given crisp set in terms of two subsets derived from a partition on the universe [Klir, Yuan, 1995], [Komorowski, Pawlak, Polkowski, Skowron, 1999], [Pawlak 1982], [Yao, 1996]. The concept of rough sets can be defined precisely as follows. Let *U* denotes a finite and non-empty universe, and let *R* be an equivalence relation on *U*. The equivalence relation *R* induces a partition of the universe. The partition is also referred to as the quotient set and is denoted by *U/R*. Suppose $[u]_R$ is the equivalence class in *U/R* that contains $u \in U$. A rough set approximation of a subset $A \subseteq U$ is a pair of lower and upper approximations. The lower approximation,

$$Lo(A) = \{u \in U \mid [u]_R \subseteq A\} = \bigcup \{[u]_R \in U/R \mid [u]_R \subseteq A\},$$

is union of all equivalence classes in *U/R* that are contained in *A*. The upper approximation,

$$Up(A) = \{u \in U \mid [u]_R \cap A \neq \emptyset\} = \bigcup \{[u]_R \in U/R \mid [u]_R \cap A \neq \emptyset\},$$

is union of all equivalence classes in *U/R* that overlap with *A*. Similarly, by rough set, a rough event can be described into two approximate events, namely lower and upper approximate events. Rough event might be considered as an approximation and generalization of a given crisp event. Probability of rough event is then defined as follows.

Definition 2.3.1 Let (*U*,*F*,*P*) be a probability space in which *U* is the sample space, *F* is sigma algebra of events and *P* is a probability measure over *U*. Then, a rough event of $A = [Lo(A), Up(A)] \in F^2$ is a pair of lower and upper approximation of $A \subseteq U$. The probability of rough event A is defined by an interval probability [*P*(*Lo*(*A*)), *P*(*Up*(*A*))], where *P*(*Lo*(*A*)) and *P*(*Up*(*A*)) are lower and upper probabilities, respectively.

- *lower probability:*

$$P(Lo(A)) = \sum_{\{u \in U | [u]_R \subseteq A\}} p(u) = \sum_{\cup\{[u]_R \in U/R | [u]_R \subseteq A\}} P([u]_R), \qquad (2.5)$$

- *upper probability:*

$$P(Lo(A)) = \sum_{\{u \in U | [u]_R \cap A \neq \varnothing\}} p(u) = \sum_{\cup\{[u]_R \in U/R | [u]_R \cap A \neq \varnothing\}} P([u]_R), \qquad (2.6)$$

where $p(u)$ is probability distribution function of element $u \in U$.

The definition shows that probability of rough event gives semantic formulation of interval probability. By combining with other set-theoretic operators such as ¬, ∪ and ∩, we have the following results:

(P1) $P(Lo(A)) \le P(A) \le P(Up(A))$,

(P2) $A \subseteq B \Leftrightarrow [P(Lo(A)) \le P(Lo(B)), P(Up(A) \le P(Up(B))]$,

(P3) $P(Lo(\neg A)) = 1 - P(Lo(A)), P(Up(\neg A)) = 1 - P(Up(A))$,

(P4) $P(\neg Lo(A)) = P(Up(\neg A)), P(\neg Up(A)) = P(Lo(\neg A))$,

(P5) $P(Lo(U)) = P(U) = P(Up(U)) = 1, P(Lo(\emptyset)) = P(\emptyset) = P(Up(\emptyset)) = 0$,

(P6) $P(Lo(A \cap B)) = P(Lo(A) \cap Lo(B)), P(Up(A \cap B)) \le P(Up(A) \cap Up(B))$,

(P7) $P(Lo(A \cup B)) \ge P(Lo(A)) + P(Lo(B)) - P(Lo(A \cap B))$,

(P8) $P(Up(A \cup B)) \le P(Up(A)) + P(Up(B)) - P(Up(A \cap B))$,

(P9) $P(A) \le P(Lo(Up(A))), P(A) \ge P(Up(Lo(A)))$,

(P10) $P(Lo(A)) = P(Lo(Lo(A))), P(Up(A)) = P(Up(Up(A)))$,

(P11) $P(Lo(A) \cup Lo(\neg A)) \le 1, P(Up(A) \cup Up(\neg A)) \ge 1$,

(P12) $P(Lo(A) \cap Lo(\neg A)) = 0, P(Up(A) \cap Up(\neg A)) \ge 0$.

Conditional probability of rough event might be considered in the following four combinations of formulation: For $A, B \subseteq U$, conditional probability of A given B is defined by

(1) $$P(Lo(A) \mid Lo(B)) = \frac{P(Lo(A) \cap Lo(B))}{P(Lo(B))},$$

(2) $P(Lo(A) \mid Up(B)) = \dfrac{P(Lo(A) \cap Up(B))}{P(Up(B))}$,

(3) $P(Up(A) \mid Lo(B)) = \dfrac{P(Up(A) \cap Lo(B))}{P(Lo(B))}$,

(4) $P(Up(A) \mid Up(B)) = \dfrac{P(Up(A) \cap Up(B))}{P(Up(B))}$,

Some relations are given by:

$$P(Lo(A) \cap Lo(B)) \leq P(Up(A) \cap Lo(B)) \Rightarrow P(Lo(A) \mid Lo(B)) \leq P(Up(A) \mid Lo(B))$$

$$P(Lo(A) \cap Up(B)) \leq P(Up(A) \cap Up(B)) \Rightarrow P(Lo(A) \mid Up(B)) \leq P(Up(A) \mid Up(B))$$

Similarly, they also satisfy some properties:

(i) Normalization:

a. $P(Lo(A)|Lo(B)) + P(Lo(\neg A)|Lo(B)) \leq 1$,

b. $P(Lo(A)|Up(B)) + P(Lo(\neg A)|Up(B)) \leq 1$,

c. $P(Up(A)|Lo(B)) + P(Up(\neg A)|Lo(B)) \geq 1$,

d. $P(Up(A)|Up(B)) + P(Up(\neg A)|Up(B)) \geq 1$.

(ii) Total Probability If $\{B_k \mid k \in N_n\}$ are crisp, pairwise disjoint and exhaustive events, i.e., $P(B_i \cap B_j) = 0$ for $i \neq j$ and $\cup B_k = U$, then:

a. $P(Lo(A)) \geq \sum_k P(Lo(B_k)) \cdot P(Lo(A) \mid Lo(B_k))$,

b. $P(Lo(A)) \leq \sum_k P(Up(B_k)) \cdot P(Lo(A) \mid Up(B_k))$,

c. $P(Up(A)) \geq \sum_k P(Lo(B_k)) \cdot P(Up(A) \mid Lo(B_k))$,

d. $P(Up(A)) \leq \sum_k P(Up(B_k)) \cdot P(Up(A) \mid Up(B_k))$,

(iii) Bayes Theorem:

a. $P(Lo(A) \mid Lo(B)) = \dfrac{P(Lo(B) \mid Lo(A)) \cdot P(Lo(A))}{P(Lo(B))}$.

b. $P(Lo(A) \mid Up(B)) = \dfrac{P(Up(B) \mid Lo(A)) \cdot P(Lo(A))}{P(Up(B))}$.

c. $P(Up(A) \mid Lo(B)) = \dfrac{P(Lo(B) \mid Up(A)) \cdot P(Up(A))}{P(Lo(B))}$.

d. $P(Up(A) \mid Up(B)) = \frac{P(Up(B) \mid Up(A)) \cdot P(Up(A))}{P(Up(B))}$.

Other considerable formulations of conditional probability of rough event are the following: For $A, B \subseteq U$, conditional probability of A given B can also be defined by,

(1) $P_1(A) \mid B) = \frac{P(Lo(A \cap B))}{P(Lo(B))}$,

(2) $P_2(A \mid B) = \frac{P(Lo(A \cap B))}{P(Up(B))}$,

(3) $P_3(A \mid B) = \frac{P(Up(A \cap B))}{P(Lo(B))}$,

(4) $P_4(A \mid B) = \frac{P(Up(A \cap B))}{P(Up(B))}$,

Also some relations concerning the above formulations are given by:

- $P_2(A|B) \leq P_1(A|B) \leq P_3(A|B)$,
- $P_4(A|B) \leq P_3(A|B)$,
- $P_2(A|B) \leq P_4(A|B)$,
- $P\,(Lo(A \cap B)) = P\,(Lo(A) \cap Lo(B)) \Rightarrow P_1(A|B) = P\,(Lo(A)|Lo(B))$.

They satisfy some properties of conditional probability:

(i) Normalization:

a. $P_1(A| B) + P_1(\neg A| B) \leq 1$,

b. $P_2(A| B) + P_2(\neg A| B) \leq 1$,

c. $P_3(A| B) + P_3(\neg A| B) \geq 1$,

d. $P_4(A| B) + P_4(\neg A| B) \geq 1$.

(ii) Total Probability If $\{B_k \mid k \in \mathrm{N}_n\}$ are crisp, pairwise disjoint and exhaustive events, i.e., $P\,(B_i \cap B_j) = 0$ for $i \neq j$ and $\cup B_k = U$, then:

a. $P(Lo(A)) \geq \sum_k P(Lo(B_k)) \cdot P_1(A \mid B_k)$,

b. $P(Lo(A)) \geq \sum_k P(Up(B_k)) \cdot P_2(A \mid B_k)$,

c. $P(Up(A)) \le \sum_k P(Lo(B_k)) \cdot P_3(A \mid B_k)$,

d. $P(Up(A)) \le \sum_k P(Up(B_k)) \cdot P_4(A \mid B_k)$.

(iii) Bayes Theorem:

a. $P_1(A \mid B) = \dfrac{P_1(B \mid A) \cdot P(Lo(A))}{P(Lo(B))}$,

b. $P_2(A \mid B) = \dfrac{P_2(B \mid A) \cdot P(Lo(A))}{P(Up(B))}$,

c. $P_3(A \mid B) = \dfrac{P_3(B \mid A) \cdot P(Up(A))}{P(Lo(B))}$,

d. $P_4(A \mid B) = \dfrac{P_4(B \mid A) \cdot P(Up(A))}{P(Up(B))}$.

2.4 Probability of Generalized Fuzzy-Rough Event

A generalized fuzzy rough set is an approximation of a given fuzzy set on a given fuzzy covering. Since fuzzy set generalizes crisp set and covering generalizes partition, fuzzy covering is regarded as the most generalized approximation space. Fuzzy covering might be considered as a case of *fuzzy granularity* in which similarity classes as a basis of constructing the covering are regarded as fuzzy sets. Alternatively, a fuzzy covering might be constructed and defined as follows [Intan, Mukaidono, 2002a].

Definition 2.4.1 Let $U = \{u_1, ..., u_n\}$ be an universe. A fuzzy covering of U is a family of fuzzy subsets or fuzzy classes of C, denoted by $C = \{C_1, C_2, ..., C_m\}$, which satisfies

$$\sum_{i=1}^{m} \mu_{C_i}(u_k) \ge 1, \quad \forall k \in \mathrm{N}_n \tag{2.7}$$

$$0 < \sum_{k=i}^{n} \mu_{C_i}(u_k) < n, \quad \forall i \in \mathrm{N}_m \tag{2.8}$$

where m is a positive integer and $\mu_{Ci}(u_k) \in [0, 1]$.

Given a fuzzy set *A* on fuzzy covering as defined in Definition 2.4.1, a generalized fuzzy rough set *A* is defined in the following definition.

Definition 2.4.2 Let *U* be a non-empty universe, $C = \{C_1, C_2, ..., C_m\}$ be a fuzzy covering and *A* be a given fuzzy set on *U*. $Lo(A)_m$, $Lo(A)_M$, $Up(A)_m$ and $Up(A)_M$ are defined as minimum lower, maximum lower, minimum upper and maximum upper approximate fuzzy set of *A*, respectively, as follows.

$$\mu_{Lo(A)_m}(y) = \inf_{\{i|\mu_{C_i}(y)>0\}} \inf_{\{z\in U|\mu_{C_i}(z)>0\}} \{\psi(i,z)\}, \tag{2.9}$$

$$\mu_{Lo(A)_M}(y) = \sup_{\{i|\mu_{C_i}(y)>0\}} \inf_{\{z\in U|\mu_{C_i}(z)>0\}} \{\psi(i,z)\}, \tag{2.10}$$

$$\mu_{Up(A)_m}(y) = \inf_{\{i|\mu_{C_i}(y)>0\}} \sup_{z\in U}\{\psi(i,z)\}, \tag{2.11}$$

$$\mu_{Up(A)_M}(y) = \sup_{\{i|\mu_{C_i}(y)>0\}} \sup_{z\in U}\{\psi(i,z)\}, \tag{2.12}$$

where $\psi(i,z) = \min(\mu_{C_i}(z), \mu_A(z))$, for short.

Therefore, a given fuzzy set *A* is approximated into four approximate fuzzy sets derived from a fuzzy covering defined on the universal set involved. Relationship among these approximations can be represented by a partial order as follows.

$$Lo(A)_m \subseteq Lo(A)_M \subseteq Up(A)_M, \quad Lo(A)_m \subseteq Up(A)_m \subseteq Up(A)_M, \quad Lo(A)_M \subseteq A.$$

Iterative is applied for almost all approximate fuzzy sets except for $Lo(A)_M$ as follows.

a. $Lo(A)_{m*} \subseteq \cdots \subseteq Lo(Lo(A)_m)_m \subseteq Lo(A)_m$,

b. $Up(A)_m \subseteq Up(Up(A)_m)_m \subseteq \cdots \subseteq Up(A)_{m*}$,

c. $Up(A)_M \subseteq Up(Up(A)_M)_M \subseteq \cdots \subseteq Up(A)_{M*}$,

where $Lo(A)_{m*}$, $Up(A)_{m*}$ and $Up(A)_{M*}$ are the lowest approximation of $Lo(A)_m$, the uppermost approximation of $Up(A)_m$ and the uppermost approximation of $Up(A)_M$, respectively.

By the generalized fuzzy rough set, a given fuzzy event can be approximated into four fuzzy events called *generalized fuzzy-rough event*. Probability of generalized fuzzy-rough event is then defined as follows.

Definition 2.4.3 Let (U, F, P) be a probability space in which U is the sample space, F is sigma algebra of events and P is a probability measure over U. Then, a generalized fuzzy-rough event of $\mathbf{A} = [Lo(A)_m, Lo(A)_M, Up(A)_m, Up(A)_M] \in F^4$ are fuzzy approximate events of A, where A is a given fuzzy event on U. The probability of generalized fuzzy-rough event A is defined by a quadruplet $[P(Lo(A)_m), P(Lo(A)_M), P(Up(A)_m), P(Up(A)_M)]$ as follows.

$$P(Lo(A)_m) = \sum_U \mu_{Lo(A)_m}(u) \cdot p(u), \tag{2.13}$$

$$P(Lo(A)_M) = \sum_U \mu_{Lo(A)_M}(u) \cdot p(u), \tag{2.14}$$

$$P(Up(A)_m) = \sum_U \mu_{Up(A)_m}(u) \cdot p(u), \tag{2.15}$$

$$P(Up(A)_M) = \sum_U \mu_{Up(A)_M}(u) \cdot p(u), \tag{2.16}$$

where $p(u)$ is probability distribution function of element $u \in U$.

By combining with other set-theoretic operators such as ¬, ∪ and ∩, we have the following properties:

a. $P(Lo(A)_m) \le P(Lo(A)_M) \le P(Up(A)_M)$,
b. $P(Lo(\mathrm{A})_M) \le P(A)$,
c. $P(Lo(A)_m) \le P(Up(A)_m) \le P(Up(A)_M)$,
d. $A \subseteq B \Rightarrow [P(Lo(A)_m) \le P(Lo(B)_m), P(Lo(A)_M) \le P(Lo(B)_M),$
$P(Up(A)_m) \le P(Up(B)_m), P(Up(A)_M) \le P(Up(B)_M)]$,
e. $P(Lo(U)_\lambda) \le 1$, $P(Up(U)_\lambda) \le 1$,
f. $P(Lo(\emptyset)_\lambda) = P(Up(\emptyset)_\lambda) = 0$,
g. $P(Lo(A \cap B)_\lambda) \le P(Lo(A)_\lambda \cap Lo(B)_\lambda)$,
h. $P(Up(A \cap B)_\lambda) \le P(Up(A)_\lambda \cap Up(B)_\lambda)$,
i. $P(Lo(A \cup B)_\lambda) \ge P(Lo(A)_\lambda) + P(Lo(B)_\lambda)) - P(Lo(A \cap B)_\lambda)$,
j. $P(Up(A \cup \mathrm{B})_\lambda) \le P(Up(A)_\lambda) + P(Up(\mathrm{B})_\lambda)) - P(Up(A \cap B)_\lambda)$,

k. $P(Lo(A)_{m*}) \le \cdots \le P(Lo(Lo(A)_m)_m) \le P(Lo(A)_m)$,

l. $P(Lo(A)_M) = P(Lo(Lo(A)_M)_M)$,

m. $P(Up(A)_m) \le P(Up(Up(A)_m)_m) \le \cdots \le P(Up(A)_{m*})$,

n. $P(Up(A)_M) \le P(Up(Up(A)_M)_M) \le \cdots \le P(Up(A)_{M*})$,

o. $P(Lo(A)_\lambda \cup Lo(\neg A)_\lambda)) \le 1$,

p. $P(Lo(A)_\lambda \cap Lo(\neg A)_\lambda)) \ge 0$,

q. $P(Up(A)_\lambda \cap Up(\neg A)_\lambda)) \ge 0$,

where $\lambda \in \{m, M\}$, for short.

2.5 Belief and Plausibility Measures

Belief and plausibility measures are mutually dual functions in evidence theory originally introduced by Glenn Shafer in 1976 [Shafer, 1976]. This work was motivated and related to lower and upper probability by Dempster in 1967 [Dempster, 1967] in which these all types of measures are subsumed into the concept of fuzzy measure proposed by Sugeno in 1977 [Sugeno, 1977]. Belief-plausibility measures can be represented by a single function, called basic probability assignment, which provides degrees of evidence to certain specific subsets of the universal set. In the special case when subsets of the universal set are disjoint and every subset represent elementary set of indiscernible space, we may consider belief measures and plausibility measures as lower approximate probability and upper approximate probability in terms of probability of rough events as proposed in [Intan, Mukaidono, 2002e], [Intan, Mukaidono, 2002g]. Here, lower and upper approximate probabilities are regarded as special case of belief and plausibility measures, respectively, as probability of elementary set is a special case of basic probability assignment. In other words, belief and plausibility measures are based on crisp-granularity in terms of a covering. However, lower and upper approximate probabilities are defined on crisp-granularity in terms of disjoint partition. Moreover, when every elementary set has only one element of set, every probability of elementary set will be equal to probability of an element called probability distribution function as usually used in representing probability measures. Obviously, lower and upper approximate probability of a given rough event will be reduced into a single value of probability. Belief and plausibility measures as well as lower and upper approximate probability are considered as generalization of probability measures in the presence of crisp granularity of sample space. Still there is another generalization in the case that membership degree of every element of sample space in representing an event might be regarded from 0 to 1. It provides probability measures of fuzzy events as proposed by Zadeh in 1968 [Zadeh, 1968]. It may then provide a more generalized probability measures in the presence of fuzzy-granularity of sample space and by given a fuzzy event called probability measures of generalized fuzzy-rough events as proposed in previous section [Intan, Mukaidono, 2002f], [Intan, Mukaidono, 2002k]. Belief and plausibility measures can be represented by a single function called basic probability assignment as defined by the following [Inuiguchi, Tanino, 2001]:

Definition 2.5.1 For U be a given universal sample space and P(U) be power set of U,

$$m : \mathrm{P}(U) \rightarrow [0, 1] \tag{2.17}$$

such that $m(\emptyset) = 0$ and $\sum_{E \in \mathcal{P}(U)} m(E) = 1$, where $m(E)$ expresses the degree of evidence supporting the claim that a specific element of U belongs to the set E but not to any special subset of E.

There are three properties considering the definition of basic probability assignment.

1. It is not required that $m(U) = 1$.
2. It is not required that $E_1 \subset E_2 \Rightarrow m(E_1) \leq m(E_2)$.
3. There is no relationship between $m(E)$ and $m(\neg E)$.

Every $E \in$ P(U) is called a *focal element* iff $m(E) > 0$. Focal elements may take overlap one to each other. Belief and Plausibility measures are then defined by the following equations. For $A \in$ P(U),

$$Bel(A) = \sum\nolimits_{E \subseteq A} m(E), \tag{2.18}$$

$$Pl(A) = \sum\nolimits_{E \cap A \neq \emptyset} m(E). \tag{2.19}$$

It can be proved that for all $A \in (P)(U)$, $Bel(A) \leq Pl(A)$. Also, it can be verified that belief and plausibility measures are mutually dual functions, where

$$Pl(A) = 1 - Bel(\neg A).$$

Similarly,

$$Bel(A) = 1 - Pl(\neg A).$$

Since belief and plausibility measures are defined on a covering, some properties of lower and upper approximate probability are not satisfied such as for instance iterative properties of lower and upper approximate probabilities in (P9) and (P10). Let consider,

$$Pl^{-1}(A) = \bigcup_{E \in \mathcal{P}(U), E \cap A \neq \emptyset} E \quad \text{and} \quad Bel^{-1}(A) = \bigcup_{E \in \mathcal{P}(U), E \subseteq A} E$$

where $Pl(A)$ and $Pl^{-1}(A)$ correspond to $P(Up(A))$ and $Up(A)$, respectively. Similarly, $Bel(A)$ and Bel^{-1}(A) correspond to $P(Lo(A))$ and $Lo(A)$, respectively. Hence, property of $P(Up(A)) = P(Up(Up(A)))$ in (P10) can be represented as Pl(A) = $Pl(Pl^{-1}(A))$ by using expression of plausibility measures. It can be easily proved that the property is not satisfied instead $Pl(A) \leq Pl(Pl^{-1}(A))$. Also, $P(A) \geq Pl(Bel^{-1}(A))$ in the relation to $P(A) \geq P(Up(Lo(A)))$, in property (P9) cannot be verified.

When every elementary set has only one element, the probability of elementary set is equal to probability of the element represented by a function called probability distribution function, $p : U \rightarrow [0,1]$, which is defined on set U as usually used in probability measures. Here, lower and upper approximate probabilities fuse into a single value of probability in which probability satisfies additivity axiom as an intersection area between supperadditive property (P7) of lower approximate probability and subadditive property (P8) of upper approximate probability.

2.6 Conclusion

The relationship between probability and fuzziness was simply discussed based on the process of perception. Probability and fuzziness work in different areas of uncertainty; hence probability theory by itself is not sufficient for dealing with uncertainty in the real-world application. Instead, probability and fuzziness must be regarded as a complementary tool providing probability of fuzzy event in which fuzzy event was represented by fuzzy set. Fuzzy event was considered as a generalization of crisp event as well as fuzzy set generalizes crisp set. Similarly, rough set, as another generalization of crisp set, was used to represent rough event. Probability of rough event was proposed. Conditional probability of fuzzy event as well as rough event and their some properties were examined. A more generalized fuzzy rough set is proposed as an approximation of a given fuzzy set on a given fuzzy covering. Therefore, by using the generalized fuzzy rough set, a generalized fuzzy-rough event was considered as the most generalization of fuzzy and rough event in terms of their definition by using probability distribution function ($p(u)$). Probability of the generalized fuzzy-rough event was proposed along with its properties.

We may then summarize their relation by the following figure.

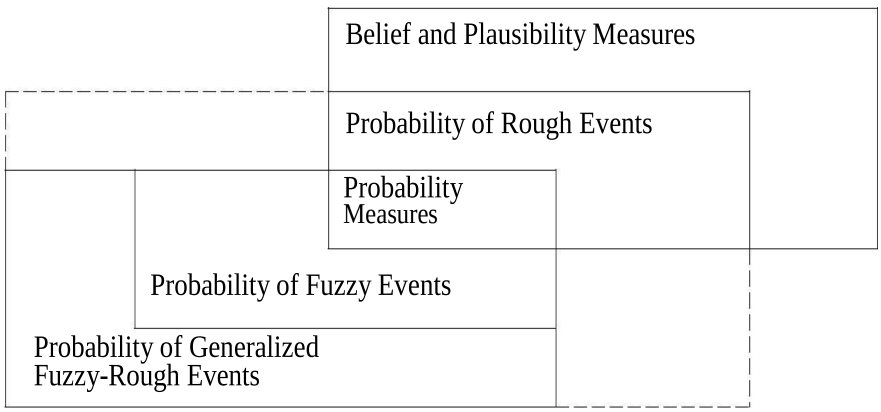

Figure 2.1 Generalization based on Crisp-Granularity and Membership Function

Chapter 3

Generalization of Rough Sets and its Applications in Information System

3.1 Introduction

Rough set theory, proposed by Pawlak in 1982, plays essential roles in many applications of Data Mining and Knowledge Discovery. The theory offers mathematical tools to discover hidden patterns in data, recognize partial or total dependencies in data bases, remove redundant data, and others [Komorowski, Pawlak, Polkowski, Skowron, 1999]. Rough set theory generalizes classical set theory by allowing an alternative to formulate sets with imprecise boundaries. A rough set is basically an approximate representation of a given crisp set in terms of two subsets derived from a crisp partition defined on the universal set involved [Klir, Yuan 1995]. The two subsets are called a lower approximation and an upper approximation. In a partition, an element belongs to the only one equivalence class and two distinct equivalence classes are disjoint. Formally, the concept of rough sets may be defined precisely, let U denotes a non-empty universal set, and let R be an equivalence relation on U. The partition of the universe is referred to as the quotient set and is denoted by U/R, where $[x]_R$ denotes the equivalence class in U/R that contains $x \in U$. A rough set of subset $A \subseteq U$ may be represented by a pair of lower and upper approximation. The lower approximation,

$$Lo(A) = \{x \in U \,|\, [x]_R \subseteq A\},$$

$$= \{[x]_R \in U/R \,|\, [x]_R \subseteq A\}, \quad (3.1)$$

is the union of all equivalence classes in *U/R* that are contained in *A*. The upper approximation,

$$Up(A) = \{x \in U \,|\, [x]_R \cap A = \emptyset\},$$

$$= \{[x]_R \in U/R \,|\, [x]_R \cap A = \emptyset\}, \quad (3.2)$$

is the union of all equivalence classes in *U/R* that overlap with *A*. Moreover, rough membership functions of an element in the presence of a subset $A \subseteq U$ is defined as the following [Pawlak, Skowron, 1994]:

$$\mu_A(x) = \frac{|[x]_R \cap A|}{|[x]_R|}, \quad (3.3)$$

where |.| denotes the cardinality of a set.

However, as pointed out in [Slowinski, Vanderpooten, 2000], even it is easy to analyze, the rough set theory built on a partition induced by equivalence relations may not provide a realistic and applicable model as equivalence relation, because of their properties of symmetry and transitivity, may not provide a realistic view of relationships between elements in real-world application. Here, covering of the universe as an alternative to provide a more realistic model of rough sets was introduced. A covering of the universe [Yao, Zhang, 2000], $C = \{C_1, ..., C_n\}$, is a family of subset of non-empty universe *U* such that $U = \{C_i \mid i = 1, ..., n\}$. Two distinct sets in *C* may have a non-empty overlap. An arbitrarily element $x \in U$ may belong to more than one set in *C*. The family $C(x) = \{C_i \in C | x \in C_i\}$ consists of sets in *C* containing *x*. The sets in *C*(*x*) may describe *different types* or various degrees of similarity between elements of *U*. In this case, relationship between elements in the set *C*(*x*) is still unclear.

In [Intan, Mukaidono, Yao, 2001a], a notion of *a weak fuzzy similarity relation*, a generalization of fuzzy similarity relation (fuzzy equivalence relation), was introduced in order to provide a more realistic relation in representing relationships between two elements in which the properties of symmetry and transitivity are no longer hold. It can be proved that the relationships between elements in the real-world application are neither necessarily symmetric nor transitive [Intan, Mukaidono, 2000d], [Slowinski, Vanderpooten, 2000], [Tversky, 1977]. (Fuzzy) conditional probability relation was then introduced as a special type (concrete example) of weak fuzzy similarity relation.

The objectives of this chapter are to extend and generalize the classical concept of rough set by coverings of the universe induced by (fuzzy) conditional probability relations. Considering the rough sets approximation, lower and upper approximations are introduced in the presence of α-covering of the universe. In this case, α determines degree of similarity relationships between elements in covering. The generalized concept of rough approximations is introduced and defined based on α-coverings of the universe into two interpretations, *element-oriented generalization* and *similarity-class-oriented generalization*. The generalized concept of rough approximations is regarded as a kind of *fuzzy rough sets*[1] approximation. A more generalized fuzzy rough set approximation of a given fuzzy set is proposed and discussed as an alternative to provide interval-valued fuzzy sets from information system. Also, a generalized concept of rough membership function is defined into three values: minimum, maximum and average. Their properties are examined.

Finally, by extending the concept of α-coverings of the universe, some applications related to knowledge discovery and data mining, were proposed and discussed with the purpose of determining redundant objects and recognizing partial and total dependency of domain attributes. The concept of redundant object is very important in order to reduce the number of decision rules in terms of decision table. The concept of dependency of domain attributes can be extended to define fuzzy functional dependency (FFD) as proposed in [Intan, Mukaidono 2000a], [Intan, Mukaidono, 2000c] in terms of fuzzy relational database.

Inference rules that are similar to Armstrong's Axioms for the FFD are both sound and complete.

3.2 Conditional Probability Relations

As proposed in [Intan, Mukaidono 2000a], [Intan, Mukaidono, 2000d], our concept of *conditional probability relations* starts from definition of an interesting mathematical relation, *weak fuzzy similarity relation* as defined in the following definition.

Definition 3.2.1 A fuzzy similarity relation is a mapping, $s: U \times U \rightarrow [0,1]$, such that for $x, y, z \in U$,

(a) Reflexivity : $s(x, x) = 1$,

(b) Symmetry : $s(x, y) = s(y, x)$,

(c) Max−min transitivity : $s(x,z) \geq \max_{y \in U}[\min[s(x,y), s(y,z)]]$.

Definition 3.2.2 A *weak fuzzy similarity relation* is a mapping, $s: U \times U \rightarrow [0, 1]$, such that for $x, y, z \in U$,

(a) Reflexivity : $s(x, x) = 1$,

(b) Conditional symmetry: if $s(x, y) > 0$ then $s(y, x) > 0$,

(c) Conditional transitivity: if $s(x, y) \geq s(y, x) > 0$ and $s(y, z) \geq s(z, y) > 0$ then $s(x, z) \geq s(z, x)$.

Definition 3.2.3 A *conditional probability relation* is a mapping, $R : U \times U \rightarrow [0,1]$, such that for $x, y \in U$,

$$R(x, y) = P(x \mid y) = P(y \rightarrow x), \tag{3.4}$$

where $R(x, y)$ means the degree y supports x or the degree y is similar to x.

It should be mentioned that symmetry and even max-min transitivity as required in fuzzy similarity relation [Zadeh, 1970] are too strong properties to represent relationships between elements in real-world application. Although, it is true to say that if "*x* is similar to *y*" then "*y* is similar to *x*", but these two statements might have different degree of similarity. In this case, degree of similarity strongly depends on which one is more general. For example, in the following two statements: "Pony is similar to horse" and "Horse is similar to pony", first statement sounds good, but the second makes much less sense. Some arguments have been proposed in [Intan, Mukaidono, 2000d], [Slowinski, Vanderpooten, 2000], [Tversky, 1977]. Hence, weak fuzzy similarity relation with its conditional symmetry and conditional transitivity properties is considered as a more realistic relation in representing relationships between elements. By definitions, fuzzy similarity relation is regarded as a special case (or type) of weak fuzzy similarity relation, and a conditional probability relation is a concrete example of weak fuzzy similarity relations. In practical application, conditional probability relations may be used as a basis of representing degree of similarity relationships between elements in the universe *U*. In the definition of conditional probability relations, the probability values may be estimated based on the semantic relationships between elements by using the epistemological or subjective view of probability theory. Relationship between *x* and *y* in conditional probability relations can be illustrated by using binary information table, where *x* and *y* are simply assumed as objects and each object is a subset of features as shown in the following table. When objects in *U* are represented by sets of features or attributes as in the case of binary information tables, we have a simple procedure for estimating the conditional probability relation. More specifically, we have:

$$R(x, y) = P(x \mid y) = \frac{|x \cap y|}{|y|}, \tag{3.5}$$

where $|\cdot|$ denotes the cardinality of a set.

Table 3.1 Binary Information Table

Obj.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
$O1$	0	0	1	0	1	0	0	0
$O2$	1	1	0	1	0	0	1	0
$O3$	0	0	1	1	0	0	1	1
$O4$	0	1	0	1	0	1	0	1
$O5$	1	0	1	1	0	0	1	0
$O6$	0	0	1	0	1	0	1	0
$O7$	0	1	1	0	0	0	1	0
$O8$	1	1	0	0	0	0	1	1
$O9$	0	1	0	1	1	0	1	0
$O10$	0	1	0	0	0	1	1	0
$O11$	0	0	0	1	1	0	1	1
$O12$	1	0	0	0	1	0	0	0
$O13$	1	0	1	0	1	0	1	0
$O14$	1	0	0	0	0	1	1	0
$O15$	0	0	1	0	1	0	1	1
$O16$	0	0	0	1	0	0	1	1
$O17$	0	1	0	1	1	0	0	1
$O18$	1	0	0	1	0	0	1	0
$O19$	0	0	1	0	1	1	0	1
$O20$	1	0	0	1	0	1	0	0

Consider the binary information table given by Table 3.1, where the set of objects, $U = \{O_1, O_2, ..., O_{20}\}$, is described by a set of eight attribute, $A = \{a_1, a_2, . . . , a_8\}$. As shown in Table 3.1, $O_1 = \{a_3, a_5\}$, $O_2 = \{a_1, a_2, a_4, a_7\}$, and $O_3 = \{a_3, a_4, a_7, a_8\}$. Therefore, we have:

$$R(O_1, O_2) = 0, \quad R(O_1, O_3) = 1/4, \quad R(O_2, O_3) = 2/4,$$
$$R(O_2, O_1) = 0, \quad R(O_3, O_1) = 1/2, \quad R(O_3, O_2) = 2/4.$$

The degree of similarity two objects can be calculated by a conditional probability relation on fuzzy sets [Intan, Mukaidono, 2000a], [Intan, Mukaidono, 2000d]. In this case, $|x| = \sum_{a \in At} \mu_x(a)$, where μ_x is membership function of *x* over a set of attribute *At*, and intersection is defined by minimum in order to obtain property of reflexivity, even there are some operations of t-norm that might be used.

Definition 3.2.4 Let μ_x and μ_y be two fuzzy sets over a set of attribute *At* for two elements *x* and *y* of a universe of objects *U*. A *fuzzy conditional probability relation* is defined by:

$$R(x,y) = \frac{\sum_{a \in At} \min[\mu_x(a), \mu_y(a)]}{\sum_{a \in At} \mu_y(a)} \quad (3.6)$$

For example, two fuzzy sets, *Warm(W)* and *Rather-Hot(RH)*, are considered as two elements or objects over {24, 26, ..., 36} in degrees Celsius as shown in Table 3.2.

Table 3.2 Fuzzy Information Table of Temperature

Object	24	26	28	30	32	34	36
W	0.2	0.5	1	1	0.5	0.2	0
RH	0	0	0	0.5	1	1	0.5

Degree of similarity relationship between W and RH is calculated by:

$$R(W,RH) = \frac{\min(1,0.5) + \min(0.5,1) + \min(0.2,1)}{0.5+1+1+0.5} = \frac{1.2}{3}$$

$$R(RH,W) = \frac{\min(1,0.5) + \min(0.5,1) + \min(0.2,1)}{0.2+0.5+1+1+0.5+0.2} = \frac{1.2}{3.4}$$

It can be easily verified that (fuzzy) conditional probability relation *R* satisfies properties of a *weak fuzzy similarity relation*. Additional properties can be found in [Intan, Mukaidono, 2000d], [Intan, Mukaidono, 2002h], [Intan, Mukaidono, 2000k]:

For $x, y, z \in U$,

(r0) $R(x, y) = R(y, x) = 1 \Leftrightarrow x = y$,

(r1) $[R(y, x) = 1, R(x, y) < 1] \Leftrightarrow x \subset y$,

(r2) $R(x, y) = R(y, x) > 0 \Rightarrow |x| = |y|$, (r3) $R(x, y) < R(y, x) \Rightarrow |x| < |y|$,

(r4) $R(x, y) > 0 \Leftrightarrow R(y, x) > 0$,

(r5) $[R(x, y) \geq R(y, x) > 0, R(y, z) \geq R(z, y) > 0] \Rightarrow R(x, z) \geq R(z, x)$.

3.3 Generalized Rough Sets Approximation

In this section, we generalize the classical concept of rough sets and concretize the concept of coverings induced by conditional probability relations. We introduce upper and lower approximations in the presence of α-coverings of the universe. First, based on conditional probability relations, we define two kinds of similarity classes of a particular element x as a basis of constructing a covering, as follows.

Definition 3.3.1 Let U be a non-empty universe, and R be a conditional probability relation on U. For any element $x \in U$, $R_s^\alpha(x)$ and $R_p^\alpha(x)$ are defined as the set that supports x and the set supported by x, respectively by as follows:

$$R_s^\alpha(x) = \{y \in U \mid R(x, y) \geq \alpha\}, \quad (3.7)$$

$$R_p^\alpha(x) = \{y \in U \mid R(y, x) \geq \alpha\}, \quad (3.8)$$

where $\alpha \in [0,1]$.

$R_s^\alpha(x)$ can also be interpreted as the set that is similar to x. On the other hand, $R_p^\alpha(x)$ can be considered as the set to which x is similar. By the reflexivity, it follows that we can construct two covering of the universe, $\{ R_s^\alpha(x) \mid x \in U\}$ and $\{ R_p^\alpha(x) \mid x \in U\}$.

Formally, based on the similarity class of x in Definition 3.3.1, the lower and upper approximation operators can be defined into two interpretations of formulation as follows.

Definition 3.3.2 For a subset $A \subseteq U$, we define two pairs of generalized rough set approximations:

(i) *element-oriented generalization:*

$$Lo(A)_e^{\alpha} = \{x \in U \mid R_s^{\alpha}(x) \subseteq A\}, \tag{3.9}$$

$$Up(A)_e^{\alpha} = \{x \in U \mid R_s^{\alpha}(x) \cap A \neq \emptyset\}, \tag{3.10}$$

(ii) *similarity-class-oriented generalization:*

$$Lo(A)_c^{\alpha} = \bigcup \{R_s^{\alpha}(x) \mid R_s^{\alpha}(x) \subseteq A, x \in U\}, \tag{3.11}$$

$$Up(A)_c^{\alpha} = \bigcup \{R_s^{\alpha}(x) \mid R_s^{\alpha}(x) \cap A \neq \emptyset, x \in U\}, \tag{3.12}$$

In Definition 3.3.2 (i), the lower approximation consists of those elements in *U* whose similarity classes are contained in *A*. The upper approximation consists of those elements whose similarity classes overlap with *A*. In Definition 3.3.2 (ii), the lower approximation is the union of all similarity classes that are contained in *A*. The upper approximation is the union of all similarity classes that overlap with *A*. Relationships among these approximations can be represented by:

$$Lo(A)_e^{\alpha} \subseteq Lo(A)_c^{\alpha} \subseteq A \subseteq Up(A)_e^{\alpha} \subseteq Up(A)_c^{\alpha}.$$

The difference between lower and upper approximations is the boundary region with respect to *A*:

$$Bnd(A)_e^{\alpha} = Up(A)_e^{\alpha} - Lo(A)_e^{\alpha}, \tag{3.13}$$

$$Bnd(A)_c^{\alpha} = Up(A)_c^{\alpha} - Lo(A)_c^{\alpha}, \tag{3.14}$$

Similarly, one can define rough set approximations based on the covering $\{R_p^{\alpha}(x) \mid x \in U\}$. The pair $(Lo(A)_e^{\alpha}, Up(A)_e^{\alpha})$ may be interpreted as a pair of set-theoretic operators on subset of the universe. It is referred to as rough set approximation operators [Yao, Zhang, 2000]. By combining with other set-theoretic operators such as ¬, ∪, and ∩, we have the following results:

(re1) $Lo(A)_e^{\alpha} = \neg Up(\neg A)_e^{\alpha}$, (re2) $Up(A)_e^{\alpha} = \neg Lo(\neg A)_e^{\alpha}$,

(re3) $Lo(A)_e^{\alpha} \subseteq A \subseteq Up(A)_e^{\alpha}$,

(re4) $Lo(\emptyset)_e^{\alpha} = Up(\emptyset)_e^{\alpha} = \emptyset$,

(re5) $Lo(U)_e^{\alpha} = Up(U)_e^{\alpha} = U$,

(re6) $Lo(A \cap B)_e^{\alpha} = Lo(A)_e^{\alpha} \cap Lo(B)_e^{\alpha}$,

(re7) $Up(A \cap B)_e^{\alpha} \subseteq Up(A)_e^{\alpha} \cap Up(B)_e^{\alpha}$,

(re8) $Lo(A \cup B)_e^{\alpha} \supseteq Lo(A)_e^{\alpha} \cup Lo(B)_e^{\alpha}$,

(re9) $Up(A \cup B)_e^{\alpha} = Up(A)_e^{\alpha} \cup Up(B)_e^{\alpha}$,

(re10) $A \neq \emptyset \Rightarrow Up(A)_e^{0} = U$,

(re11) $A \subset U \Rightarrow Lo(A)_e^{0} = \emptyset$,

(re12) $\alpha \leq \beta \Rightarrow [Up(A)_e^{\beta} \subseteq Up(A)_e^{\alpha}, Lo(A)_e^{\alpha} \subseteq Lo(A)_e^{\beta}]$,

(re13) $A \subseteq B \Rightarrow [Up(A)_e^{\alpha} \subseteq Up(B)_e^{\alpha}, Lo(A)_e^{\alpha} \subseteq Lo(B)_e^{\alpha}]$.

Property (re1) and (re2) shows that lower and upper approximations are dual operators with respect to set complement $\neg$. Property (re3) shows that the two operators provide a range in which lies the given set. Properties (re4) and (re5) provide two boundary conditions, the minimum element $\emptyset$, the maximum element U; where the two operators meet at the two extreme points of 2^U. Properties (re6) and (re7) may be considered as weak distributive and distributive of the lower approximation and the upper approximation operators over set intersection and union, respectively. As shown in property (re10), if $\alpha = 0$ then all elements in U belong to one similarity class which is equal to U. Therefore, upper approximation of a non-empty set A is equal to U. On the other hand, property (re11) shows that lower approximation of A is equal to $\emptyset$ for $A \subset U$. Property (re12) shows relationships between α and the two approximation operators. Here, whenever α is getting larger, the upper approximation is getting smaller. On the other hand, whenever α is getting larger, the lower approximation is also bigger. In the other word, as α is getting larger, the process of approximation is more precise and nearer to the original set. Property (re13) indicates the consistency of inclusive sets; where if $A \subseteq B$, then both of their two approximation operators show the same characteristics.

Similarly, lower and upper approximations in Definition 3.3.2 (ii) satisfy some properties such as:

(rc1) $Lo(A)_e^\alpha = \neg Up(\neg A)_e^\alpha$, $Lo(A)_c^\alpha \subseteq A \subseteq Up(A)_c^\alpha$,

(rc2) $Up(A)_e^\alpha = \neg Lo(\neg A)_e^\alpha$, $Lo(\varnothing)_c^\alpha = Up(\varnothing)_c^\alpha = \varnothing$,

(rc3) $Lo(U)_c^\alpha = Up(U)_c^\alpha = U$,

(rc4) $Lo(A \cap B)_c^\alpha \subseteq Lo(A)_c^\alpha \cap Lo(B)_c^\alpha$,

(rc5) $Up(A \cap B)_c^\alpha \subseteq Up(A)_c^\alpha \cap Up(B)_c^\alpha$,

(rc6) $Lo(A \cup B)_c^\alpha \supseteq Lo(A)_c^\alpha \cup Lo(B)_c^\alpha$,

(rc7) $Up(A \cup B)_c^\alpha = Up(A)_c^\alpha \cup Up(B)_c^\alpha$,

(rc8) $Lo(A)_c^\alpha = Lo(Lo(A)_c^\alpha)_c^\alpha$,

(rc9) $Up(A)_c^\alpha = Up(Up(A)_c^\alpha)_c^\alpha$,

(rc10) $A \neq \varnothing \Rightarrow Up(A)_c^0 = U$,

(rc11) $A \subset U \Rightarrow Lo(A)_c^0 = \varnothing$,

(rc12) $\alpha \leq \beta \Rightarrow Up(A)_c^\beta \subseteq Up(A)_c^\alpha$,

(rc13) $A \subseteq B \Rightarrow [Up(A)_c^\alpha \subseteq Up(B)_c^\alpha, Lo(A)_c^\alpha \subseteq Lo(B)_c^\alpha]$.

Property of dual operators is no longer satisfied in Definition 3.3.2 (ii). On the other hand, property (rc8) indicates that iterative operation is not applied in the lower approximation operator. The above properties show almost the same properties which are also satisfied in classical concept of rough sets, except that they have additional parameter α and its relation to both of operators, the lower approximation and the upper approximation, as shown in properties (re10,rc10), (re11,rc11), and (re12, rc12). In fact, a covering is a generalization of a partition, so that there are some properties which are no longer satisfied.

Also one may define other interpretation of pair approximation operators based on intersection of the complements of elements as well as the complements of similarity classes [Inuiguchi, Tanino, 2001] as shown in the following equations.

(i) *element-oriented generalization:*

$$Lo(A)_{e1}^{\alpha} = \bigcap \{U - \{x\} \mid R_s^{\alpha}(x) \cap (U - A) \neq \varnothing\}, \tag{3.15}$$

$$Up(A)_{e1}^{\alpha} = \bigcap \{U - \{x\} \mid R_s^{\alpha}(x) \cap A = \varnothing\}, \tag{3.16}$$

(ii) *similarity-class-oriented generalization:*

$$Lo(A)_{c1}^{\alpha} = \bigcap \{U - R_s^{\alpha}(x) \mid R_s^{\alpha}(x) \cap (U - A) \neq \varnothing\}, \tag{3.17}$$

$$Up(A)_{c1}^{\alpha} = \bigcap \{U - R_s^{\alpha}(x) \mid R_s^{\alpha}(x) \cap A = \varnothing\}, \tag{3.18}$$

Related to the approximation operators as defined in Definition 3.3.2 (based on union of both elements and similarity classes), we can prove

$$Lo(A)_{e}^{\alpha} = Lo(A)_{e1}^{\alpha} \subseteq A, \qquad Lo(A)_{c1}^{\alpha} \subseteq Lo(A)_{c}^{\alpha} \subseteq A,$$

$$A \subseteq Up(A)_{e}^{\alpha} = Up(A)_{e1}^{\alpha}, \qquad A \subseteq Up(A)_{c1}^{\alpha} \subseteq Up(A)_{c}^{\alpha}.$$

In element-oriented generalization, lower and upper approximation operators based on both union and intersection are exactly the same. However, in similarity class-oriented generalization, $Lo(A)_{c}^{\alpha}$ is a better lower approximation than $Lo(A)_{c1}^{\alpha}$, but $Up(A)_{c1}^{\alpha}$ is a better upper approximation than $Up(A)_{c}^{\alpha}$. Here, we cannot verify relation between $Lo(A)_{e1}^{\alpha}$ and $Lo(A)_{c1}^{\alpha}$ as well as $Up(A)_{e1}^{\alpha}$ and $Up(A)_{c1}^{\alpha}$. Similarly, one may use $R_p^{\alpha}(x)$ to define approximation operators as given in (3.15)-(3.18), and verify their properties.

3.4 Generalized Fuzzy Rough Sets

We may consider the rough set approximation in the previous definition as a kind of *fuzzy rough set* which was introduced first time in [Dubois, Prade 1990].

Covering of the universe in Definition 3.3.1 as a generalization of disjoint partition is considered as a crisp covering. Both crisp covering and disjoint partition are regarded as *crisp granularity*. Here, crisp covering can be generalized to fuzzy covering. In this case, crisp covering can be constructed by applying α-level set of fuzzy covering. Fuzzy covering might be considered as a case of *fuzzy granularity* in which similarity classes as basis of constructing the covering are regarded as fuzzy sets and defined as follows.

Definition 3.4.1 Let U be a non-empty universe, and R be a (fuzzy) conditional probability relation on U. For any element $x \in U$, $R_s(x)$ and $R_p(x)$ are regarded as fuzzy sets and defined as the set that supports x and the set supported by x, respectively by:

$$\mu_{R_s(x)}(y) = R(x, y), \quad y \in U, \tag{3.19}$$

$$\mu_{R_p(x)}(y) = R(y, x), \quad y \in U, \tag{3.20}$$

where $\mu_{R_s(x)}(y)$ and $\mu_{R_p(x)}(y)$ are grades of membership of y in $R_s(x)$ and $R_p(x)$, *respectively.*

Now, when we consider a given set A be a fuzzy set on U instead of a crisp set and covering of the universe be a fuzzy covering (Definition 3.4.1) instead of a crisp covering (Definition 3.3.1), we need to define a more generalized fuzzy rough set approximation of a given fuzzy set as shown in the following definition.

Definition 3.4.2 Let U be a non-empty universe, and A be a given fuzzy set on U,

(i) *element-oriented generalization:*

$$\mu_{Lo(A)_{e2}}(x) = \inf_{\{y \in U | \mu_{R_s(x)}(y) > 0\}} \{\min[\mu_{R_s(x)}(y), \mu_A(y)]\}, \tag{3.21}$$

$$\mu_{Up(A)_{e2}}(x) = \sup_{\{y \in U | \mu_{R_s(x)}(y) > 0\}} \{\min[\mu_{R_s(x)}(y), \mu_A(y)]\}. \tag{3.22}$$

(ii) *similarity-class-oriented generalization, for y ∈U:*

$$\mu_{Lo(A)^m_{c2}}(y) = \inf_{\{x\in U|\mu_{R_s(x)}(y)>0\}}\{\inf_{\{z\in U|\mu_{R_s(x)}(z)>0\}}\{\min[\mu_{R_s(x)}(z),\mu_A(z)]\}\}, \quad (3.23)$$

$$\mu_{Lo(A)^M_{c2}}(y) = \sup_{\{x\in U|\mu_{R_s(x)}(y)>0\}}\{\inf_{\{z\in U|\mu_{R_s(x)}(z)>0\}}\{\min[\mu_{R_s(x)}(z),\mu_A(z)]\}\}, \quad (3.24)$$

$$\mu_{Up(A)^m_{c2}}(y) = \inf_{\{x\in U|\mu_{R_s(x)}(y)>0\}}\{\sup_{z\in U}\{\min[\mu_{R_s(x)}(z),\mu_A(z)]\}\}, \quad (3.25)$$

$$\mu_{Up(A)^M_{c2}}(y) = \sup_{\{x\in U|\mu_{R_s(x)}(y)>0\}}\{\sup_{z\in U}\{\min[\mu_{R_s(x)}(z),\mu_A(z)]\}\}. \quad (3.26)$$

where $\mu_{Lo(A)_{e2}}(x)$ and $\mu_{Up(A)_{e2}}(x)$ are grades of membership of x in $Lo(A)_{e2}$ and $Up(A)_{e2}$, respectively. Similarly, $\mu_{Lo(A)^*_{c2}}(y)$ and $\mu_{Up(A)^*_{c2}}(y)$ are grades of membership of y in $Lo(A)^*_{c2}$ and $Up(A)^*_{c2}$, respectively (Note: $* \in \{m, M\}$).

Since $\mu_{Rs}(x)(y) = \mu_{Rp}(y)(x)$ as shown in Definition 3.4.1, we may represent Definition 3.4.2 by using R_p as follows:

(i) *element-oriented generalization:*

$$\mu_{Lo(A)_{e2}}(x) = \inf_{\{y\in U|\mu_{R_p(y)}(x)>0\}}\{\min[\mu_{R_p(y)}(x),\mu_A(y)]\}, \quad (3.27)$$

$$\mu_{Up(A)_{e2}}(x) = \sup_{\{y\in U|\mu_{R_p(y)}(x)>0\}}\{\min[\mu_{R_p(y)}(x),\mu_A(y)]\}. \quad (3.28)$$

(ii) *similarity-class-oriented generalization, for y ∈U:*

$$\mu_{Lo(A)^m_{c2}}(y) = \inf_{\{x\in U|\mu_{R_p(y)}(x)>0\}}\{\inf_{\{z\in U|\mu_{R_p(z)}(x)>0\}}\{\min[\mu_{R_p(z)}(x),\mu_A(z)]\}\}, \quad (3.29)$$

$$\mu_{Lo(A)^M_{c2}}(y) = \sup_{\{x\in U|\mu_{R_p(y)}(x)>0\}}\{\inf_{\{z\in U|\mu_{R_p(z)}(x)>0\}}\{\min[\mu_{R_p(z)}(x),\mu_A(z)]\}\}, \quad (3.30)$$

$$\mu_{Up(A)^m_{c2}}(y) = \inf_{\{x\in U|\mu_{R_p(y)}(x)>0\}}\{\sup_{z\in U}\{\min[\mu_{R_p(z)}(x),\mu_A(z)]\}\}, \quad (3.31)$$

$$\mu_{Up(A)^M_{c2}}(y) = \sup_{\{x\in U|\mu_{R_p(y)}(x)>0\}}\{\sup_{z\in U}\{\min[\mu_{R_p(z)}(x),\mu_A(z)]\}\}. \quad (3.32)$$

Obviously, $Lo(A)_{e2}$ and $Up(A)_{e2}$ as well as $Lo(A)^*_{c2}$ and $Up(A)^*_{c2}$ are considered as fuzzy sets, where we have, $\forall y \in U$,

$$\mu_{Lo(A)_{e2}}(y) \le \mu_A(y) \le \mu_{Up(A)_{e2}}(y),$$

$$\mu_{Lo(A)^m_{c2}}(y) \le \mu_{Lo(A)^M_{c2}}(y) \le \mu_A(y) \le \mu_{Up(A)^m_{c2}}(y) \le \mu_{Up(A)^M_{c2}}(y),$$

Moreover, relation between element-oriented generalization and similarity-class-oriented generalization is represented by,

$$\mu_{Lo(A)_{e2}}(y) \le \mu_{Lo(A)_{c2}^{M}}(y) \text{ and } \mu_{Up(A)_{e2}}(y) \le \mu_{Up(A)_{c2}^{M}}(y),$$

where relation between $\mu_{Lo(A)_{e2}}(y)$ and $\mu_{Lo(A)_{c2}^{m}}(y)$ as well as relation between $\mu_{Up(A)_{e2}}(y)$ and $\mu_{Up(A)_{c2}^{m}}(y)$ cannot be inquired.

Also, we consider the pairs of $\langle \mu_{Lo(A)_{e2}}(y), \mu_{Up(A)_{e2}}(y) \rangle$ and $\langle \mu_{Lo(A)_{c2}^{*}}(y), \mu_{Up(A)_{c2}^{*}}(y) \rangle$ as lower and upper membership function of y in A. Lower and upper membership functions are the bounds of an interval value characterized by interval valued type-2 fuzziness [Turksen, 2001]. In this case, Definition 3.4.2 shows an alternative to obtain interval-valued fuzzy set from information system via generalized fuzzy rough sets approximation of fuzzy set. Let A be defined as an interval-valued fuzzy set given an ordinary fuzzy set A. For $y \in U$,

$$\mu_{\mathrm{A}}(y) = \langle \mu_{Lo(A)_{e2}}(y), \mu_{Up(A)_{e2}}(y) \rangle \text{ or}$$

$$\mu_{\mathrm{A}}(y) = \langle \mu_{Lo(A)_{c2}^{*}}(y), \mu_{Up(A)_{c2}^{*}}(y) \rangle,$$

where the pair of $\langle \mu_{Lo(A)_{c2}^{*}}(y), \mu_{Up(A)_{c2}^{*}}(y) \rangle$ can be represented by either the pair of $\langle \mu_{Lo(A)_{c2}^{m}}(y), \mu_{Up(A)_{c2}^{m}}(y) \rangle$ as well as $\langle \mu_{Lo(A)_{c2}^{M}}(y), \mu_{Up(A)_{c2}^{M}}(y) \rangle$.

3.5 Generalized Rough Membership Functions

As pointed out in [Yao, 1996], there are at least two views which can be used to interpret the rough set theory, operator-oriented view and set-oriented view. In this chapter, the operator-oriented view has been proposed in the previous section providing the lower approximation and the upper approximation operators in the presence of α-coverings. In this section, we provide the set-oriented view based on the notion of rough membership functions. In this case, rough membership functions of an element will be expressed into three values: minimum, maximum and average depending on similarity classes that cover the element as shown in the following definition.

Definition 3.5.1 Let $A \subseteq U$ be a crisp set, where U is a non-empty universe, and let R be a conditional probability relation on U. $R_s^{\alpha}(x) = \{y \in U \mid R(x,y) \geq \alpha\}$, denotes similarity class of x with α-cut, where $\alpha \in [0,1]$. $\mu_A^m(y)^{\alpha}$, $\mu_A^M(y)^{\alpha}$ and $\mu_A^*(y)^{\alpha}$ are defined as minimum, maximum and average rough membership functions of y with α-cut in the presence of set A, respectively as follows:

$$\mu_A^m(y)^{\alpha} = \min\left\{\frac{|R_s^{\alpha}(x) \cap A|}{|R_s^{\alpha}(x)|} \mid x \in U, y \in R_s^{\alpha}(x)\right\}, \tag{3.33}$$

$$\mu_A^M(y)^{\alpha} = \max\left\{\frac{|R_s^{\alpha}(x) \cap A|}{|R_s^{\alpha}(x)|} \mid x \in U, y \in R_s^{\alpha}(x)\right\}, \tag{3.34}$$

$$\mu_A^*(y)^{\alpha} = \mathrm{avg}\left\{\frac{|R_s^{\alpha}(x) \cap A|}{|R_s^{\alpha}(x)|} \mid x \in U, y \in R_s^{\alpha}(x)\right\}. \tag{3.35}$$

The above definition generalizes the concept of rough membership functions [Pawlak, Skowron, 1994] and concretizes definition of generalized rough membership functions based on a covering of the universe [Yao, Zhang, 2000]. In this case, the minimum, the maximum and the average equations may be assumed as the most pessimistic, the most optimistic and the balanced view in defining rough membership functions. The minimum rough membership function of y is determined by a set, $R_s^{\alpha}(x)$ to which y belongs, which has the smallest overlap with A compared to the cardinality of the set relatively. On the other hand, the maximum rough membership function is determined by a set, $R_s^{\alpha}(x)$ to which y belongs which has the largest overlap with A compared to the cardinality of the set relatively. The average rough membership function depends on the average of every set, $R_s^{\alpha}(x)$ to which y belongs. The relationships of the three rough membership functions can be expressed by:

$$\mu_A^m(y)^{\alpha} \leq \mu_A^*(y)^{\alpha} \leq \mu_A^M(y)^{\alpha}.$$

Moreover, the three rough membership functions may take varied values in calculation depending on the value of α. The minimum, maximum and average rough membership functions satisfy some properties as follows: for $A, B \subseteq U$ are crisp sets, where U is a non-empty universe,

(gr1) $\mu_U^m(x)^\alpha = \mu_U^*(x)^\alpha = \mu_U^M(x)^\alpha = 1$,

(gr2) $\mu_\varnothing^m(x)^\alpha = \mu_\varnothing^*(x)^\alpha = \mu_\varnothing^M(x)^\alpha = 0$,

(gr3) $[\forall R_s^\alpha(x), y \in R_s^\alpha(x) \Leftrightarrow z \in R_s^\alpha(x)] \Rightarrow [\mu_A^m(y)^\alpha = \mu_A^m(z)^\alpha, \mu_A^*(y)^\alpha = \mu_A^*(z)^\alpha,$
$\mu_A^M(y)^\alpha = \mu_A^M(z)^\alpha]$,

(gr4) $y, z \in R_s^\alpha(x) \Rightarrow [\mu_A^m(y)^\alpha \neq 0 \Rightarrow \mu_A^m(z)^\alpha \neq 0, \mu_A^m(y)^\alpha = 1 \Rightarrow \mu_A^m(z)^\alpha = 1]$,
$y \in A \Rightarrow \mu_A^m(y)^\alpha > 0$,

(gr5) $\mu_A^M(y)^\alpha = 1 \Rightarrow y \in A$,

(gr6) $A \subseteq B \Rightarrow [\mu_A^m(y)^\alpha \leq \mu_B^m(y)^\alpha, \mu_A^*(y)^\alpha \leq \mu_B^*(y)^\alpha, \mu_A^m(y)^M \leq \mu_B^M(y)^\alpha]$,

(gr7) $A \neq \varnothing \Rightarrow \mu_A^m(x)^0 = \mu_A^*(x)^0 = \mu_A^M(x)^0 = \frac{|A|}{|U|} = P(A)$.

Properties (gr1) and (gr2) show the boundaries condition of set, U and $\varnothing$, where minimum, maximum and average membership functions have the same values for all elements, 1 and 0, respectively. Properties (gr3) and (gr4) indicate that two similar elements in coverings should have similar rough membership functions. Property (gr5) and (gr6) can be used to prove the real member of a given crisp set. Property (gr7) shows that the consistency of inclusive sets should have the same characteristics in comparison between their rough membership functions. All coverings of the universe will be equal to the universe if α is equal to 0. In this case, all rough membership functions of all elements are equal to the probability of a given non-empty crisp set in the universe as shown in property (gr8).

Related to set-theoretic operators, ¬, ∩, and ∪, the rough membership functions satisfy some properties such as:

(g1) $\mu_{\neg A}^{m}(x)^{\alpha} = 1 - \mu_{A}^{M}(x)^{\alpha}$,

(g2) $\mu_{\neg A}^{M}(x)^{\alpha} = 1 - \mu_{A}^{m}(x)^{\alpha}$,

(g3) $\mu_{\neg A}^{*}(x)^{\alpha} = 1 - \mu_{A}^{*}(x)^{\alpha}$,

(g4) $\max(0, \mu_{A}^{m}(x)^{\alpha} + \mu_{B}^{m}(x)^{\alpha} - \mu_{A\cup B}^{M}(x)^{\alpha}) \le \mu_{A\cap B}^{m}(x)^{\alpha} \le \min(\mu_{A}^{m}(x)^{\alpha}, \mu_{B}^{m}(x)^{\alpha})$,

(g5) $\max(\mu_{A}^{M}(x)^{\alpha}, \mu_{B}^{M}(x)^{\alpha}) \le \mu_{A\cup B}^{M}(x)^{\alpha} \le \min(1, \mu_{A}^{M}(x)^{\alpha} + \mu_{B}^{M}(x)^{\alpha} - \mu_{A\cap B}^{m}(x)^{\alpha})$,

(g6) $\mu_{A\cup B}^{*}(x)^{\alpha} = \mu_{A}^{*}(x)^{\alpha} + \mu_{A}^{*}(x)^{\alpha} - \mu_{A\cap B}^{*}(x)^{\alpha}$.

By definition, generalized rough membership function provides four regions of $A \subseteq U$ as defined as follows:

1. Very positive region of A: $vpos(A) = \{x \in U \mid \mu_{A}^{m}(x)^{\alpha} = 1\}$,
2. Positive region of A: $pos(A) = \{x \in U \mid \mu_{A}^{m}(x)^{\alpha} > 0\}$,
3. Ambiguous region (boundary) of A: $bnd(A) = \{x \in U \mid \mu_{A}^{m}(x)^{\alpha} = 0, \mu_{A}^{M}(x)^{\alpha} > 0\}$,
4. Negative region of A: $neg(A) = \{x \in U \mid \mu_{A}^{M}(x)^{\alpha} = 0\}$.

It is necessary to denote some properties such that

- $vpos(A) \subseteq pos(A)$,
- $x \in A \Rightarrow x \in pos(A)$,
- $x \in bnd(A) \Rightarrow x \in A$,
- $pos(A) \cap bnd(A) \cap neg(A) = \emptyset$
- $pos(A) \cup bnd(A) \cup neg(A) = U$.

Also, one can defines $pos(A) - vpos(A)$ as a boundary of positive region or gives a special attention to the region in which $\mu_{A}^{M}(x)^{\alpha} = 1$ as a part of positive region.

Covering of the universe in Definition 3.3.1 as a generalization of disjoint partition proposed in classical rough set is considered as a crisp covering. Both crisp covering and disjoint partition are regarded as *crisp* *granularity*. However, when we consider a given set A be a fuzzy set on U instead of a crisp set and covering of the universe be a fuzzy covering instead of a crisp covering, we need to define a more generalized rough membership function of Definition 3.5.1. Here, fuzzy covering generalizes crisp covering. In this case, crisp covering can be constructed by applying α-level set of fuzzy covering. Fuzzy covering might be considered as a case of *fuzzy granularity* in which similarity classes as a basis of constructing the covering are regarded as fuzzy sets as defined in Definition 3.4.1.

Similarly, a more generalized rough membership function is also defined into three values: minimum, maximum and average.

Definition 3.5.2 Let A be a fuzzy set on U, where U is a non-empty universe, and let R be a (fuzzy) conditional probability relation on U. $\mu_A^m(y)$, $\mu_A^M(y)$ and $\mu_A^*(y)$ are defined as minimum, maximum and average rough membership functions of y in the presence of fuzzy set A, respectively as follows:

$$\mu_A^m(y) = \min\left\{\frac{|R_s(x)\cap A|}{|R_s(x)|} \mid x\in U, y\in R_s(x)\right\}, \qquad (3.36)$$

$$\mu_A^M(y) = \max\left\{\frac{|R_s(x)\cap A|}{|R_s(x)|} \mid x\in U, y\in R_s(x)\right\}, \qquad (3.37)$$

$$\mu_A^*(y) = \text{avg}\left\{\frac{|R_s(x)\cap A|}{|R_s(x)|} \mid x\in U, y\in R_s(x)\right\}, \qquad (3.38)$$

where

$|R_s(x)\cap A| = \sum_{z\in U}\min[\mu_{R_S(x)}(z), \mu_A(z)]$ and $|R_s(x)| = \sum_{z\in U}\mu_{R_S(x)}(z)$

.

Again, intersection is defined by minimum function in order to obtain property of reflexivity, although there are some operations of t-norm that might be used.

3.6 Illustrative Example

Let us illustrate the above concepts by using binary information table as shown in Table 3.1. Given $X \subset U$ be a crisp set of object, where:

$$X = \{O_2, O_4, O_7, O_8, O_{13}, O_{16}, O_{17}\}.$$

Given an arbitrary α-cut is equal to 0.75. Thus, we just consider to similarity classes of covering in which minimum degree of similarity concerning relationships between elements is equal to 0.75. By Definition 3.2.3, 3.2.4 and 3.3.1, we construct similarity classes of all elements in U which also represent covering of the universe by:

$R_s^{0.75}(O_1) = \{O_1\}$, $R_s^{0.75}(O_{11}) = \{O_3, O_9, O_{11}, O_{13}, O_{15}\}$,

$R_s^{0.75}(O_2) = \{O_2, O_5, O_8, O_9, O_{18}\}$, $R_s^{0.75}(O_{12}) = \{O_{12}\}$,

$R_s^{0.75}(O_3) = \{O_3, O_5, O_{11}, O_{15}, O_{16}\}$, $R_s^{0.75}(O_{13}) = \{O_1, O_5, O_6, O_{12}, O_{13}, O_{15}\}$,

$R_s^{0.75}(O_4) = \{O_4, O_{17}\}$, $R_s^{0.75}(O_{14}) = \{O_{14}\}$,

$R_s^{0.75}(O_5) = \{O_2, O_3, O_5, O_{13}\}$, $R_s^{0.75}(O_{15}) = \{O_1, O_3, O_6, O_{11}, O_{13}, O_{15}, O_{19}\}$,

$R_s^{0.75}(O_6) = \{O_1, O_6, O_{13}, O_{15}\}$, $R_s^{0.75}(O_{16}) = \{O_3, O_{11}, O_{16}\}$,

$R_s^{0.75}(O_7) = \{O_7\}$, $R_s^{0.75}(O_{17}) = \{O_4, O_9, O_{11}, O_{17}\}$,

$R_s^{0.75}(O_8) = \{O_2, O_8\}$, $R_s^{0.75}(O_{18}) = \{O_2, O_5, O_{18}\}$,

$R_s^{0.75}(O_9) = \{O_2, O_9, O_{11}, O_{17}\}$, $R_s^{0.75}(O_{19}) = \{O_1, O_{15}, O_{19}\}$,

$R_s^{0.75}(O_{10}) = \{O_{10}\}$, $R_s^{0.75}(O_{20}) = \{O_{20}\}$.

Now, we calculate all approximation operators representing approximation space based on Definition 3.3.2 and (3.9)-(3.12) as the following:

$Lo(X)_e^{0.75} = \{O_4, O_7, O_8\}$,

$Up(X)_e^{0.75} = \{O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_9, O_{11}, O_{13}, O_{15}, O_{16}, O_{17}, O_{18}\}$,

$Lo(X)_c^{0.75} = \{O_2, O_4, O_7, O_8, O_{17}\}$,

$Up(X)_c^{0.75} = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_9, O_{11}, O_{12}, O_{13}, O_{15}, O_{16}, O_{17}, O_{18}, O_{19}\}$,

$Lo(X)_{e1}^{0.75} = \{O_4, O_7, O_8\}$,

$$Up(X)_{e1}^{0.75} = \{O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_{9,} O_{11}, O_{13}, O_{15}, O_{16}, O_{17}, O_{18}\},$$

$$Lo(X)_{c1}^{0.75} = \{O_4, O_7, O_8\},$$

$$Up(X)_{c1}^{0.75} = \{O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_{9,} O_{11}, O_{13}, O_{16}, O_{17}, O_{18}\}.$$

Rough boundaries of X are the following:

$$Bnd(X)_{e}^{0.75} = \{O_2, O_3, O_5, O_6, O_9, O_{11}, O_{13}, O_{15}, O_{16}, O_{17}, O_{18}\},$$

$$Bnd(X)_{c}^{0.75} = \{O_1, O_3, O_5, O_6, O_9, O_{11}, O_{13}, O_{15}, O_{16}, O_{17}, O_{18}, O_{19}\},$$

$$Bnd(X)_{e1}^{0.75} = \{O_2, O_3, O_5, O_6, O_9, O_{11}, O_{13}, O_{15}, O_{16}, O_{17}, O_{18}\},$$

$$Bnd(X)_{c1}^{0.75} = \{O_2, O_3, O_5, O_6, O_9, O_{11}, O_{13}, O_{16}, O_{17}, O_{18}\}.$$

Next, by Definition 3.5.1, we examine and calculate minimum, maximum, and average rough membership functions of element or object O_{17}, for instance. In this case, there are three similarity classes or coverings to which O_{17} belongs. They are $R_s^{0.75}(O_4)$, $R_s^{0.75}(O_9)$ and $R_s^{0.75}(O_{17})$, where:

$$\frac{|R_s^{0.75}(O_4) \cap X|}{|R_s^{0.75}(O_4)|} = 1, \quad \frac{|R_s^{0.75}(O_9) \cap X|}{|R_s^{0.75}(O_9)|} = \frac{2}{4}, \quad \frac{|R_s^{0.75}(O_{17}) \cap X|}{|R_s^{0.75}(O_{17})|} = \frac{2}{4}.$$

Finally, by the above results, we calculate minimum, maximum and average rough membership function of O_{17} as follows:

$$\mu_A^m(O_{17})^{0.75} = min\{1,\ 2/4,\ 2/4\} = 2/4,$$

$$\mu_A^M(O_{17})^{0.75} = max\{1,\ 2/4,\ 2/4\} = 1,$$

$$\mu_A^*(O_{17})^{0.75} = \text{avg}\{1,\ 2/4,\ 2/4\} = 2/3.$$

By rough membership function of all objects, four regions of X are given as:

$vpos(X) = \{O_7\}$,
$pos(X) = \{O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_9, O_{11}, O_{13}, O_{16}, O_{17}, O_{18}\}$,
$bnd(X) = \{O_1, O_{15}, O_{19}\}$,
$neg(X) = \{O_{10}, O_{12}, O_{14}, O_{20}\}$.

Fuzzy information table can be used to generalize binary information table of Table 3.1. Here, fuzzy conditional probability relation as defined in Definition 3.2.4 is used to construct α-covering of the universe.

3.7 α-Redundancy of Objects

A data table, called *information system*, contains data about objects of interest characterized by some domain attributes. Possibly, several objects have nearly the same characteristics in a given information system. Hence, some of them may be considered as redundant objects to the others. When information systems are applied to decision systems, an object u is a non-redundant object if there is no proper u' that covers u concerning their characteristics by domain attributes. If such a u' exists, u is redundant. Simply, an example given in Table 3.3 shows a data table of *Reproduction of Animals* which is obviously proved that object u_1 is a redundant object, for it is covered by u_2 in which horse belongs to group of mammals.

Table 3.3 Reproduction of Animals

U	Description	Reproduction
u_1	Horse	Bear
u_2	Mammals	Bear
u_3	Bird	Egg

In this section, we propose a concept of determining α-redundant objects based on the concept of α-coverings of the universe. In previous section, we proposed the concept of α-coverings of the universe generalizing classical rough sets. Principally, every class in a covering corresponds to similarity class of an element or object in the universe constructed by degree of similarity, α. A redundant object is considered as object whose class is subset of class of another object in the universe. Formally, an information system is defined as a pair $I = (U, A)$, where U is a non-empty finite set of *objects* called the *universe* and A is a non-empty finite set of *domain attributes* such that $a : U \rightarrow V_a$ for every $a \in A$. The set

V_a is called the *value set of a*, where V_a may consist of precise as well as imprecise data. First, we define similarity class of object extending definition of similarity class as defined in Definition 3.3.1.

Definition 3.7.1 Let I = (U, A) be an information system, and R be a (fuzzy) conditional probability relation. For any object $u_i \in U$, $S_A^{\alpha}(u_i)$ is defined as the set that is similar to u_i by employing A to degrees of at least α in which $\alpha = (\alpha_1, ..., \alpha_n)$ corresponds to $A = (a_1, ..., a_n)$ such that:

$$S_A^{\alpha}(u_i) = \{u \in U \mid R(a_k(u_i), a_k(u)) \geq a_k, k = 1, \ldots, n\}, \qquad (3.39)$$

where $\alpha_k \in [0, 1]$. $a_k(u_i)$, $a_k(u) \in V_{ak}$ denote the restriction of the domain attribute a_k to the objects u_i and u, respectively. By Definition 3.7.1, a redundant object is determined as the following:

Definition 3.7.2 An object $u_i \in U$ is a α-redundant object in an information system $I = (U, A)$ if there is an object $u_j \in U$ whenever:

$$S_A^{\alpha}(u_i) \subseteq S_A^{\alpha}(u_j), \qquad (3.40)$$

where $S_A^{\alpha}(u_i)$ and $S_A^{\alpha}(u_j)$ are similarity classes of u_i and u_j, respectively by employing A to degrees of at least α as defined in Definition 3.7.1.

In case of crisp data are applied in classical rough sets, their degree of similarity is either 0 or 1 (0 if they are different, otherwise 1 if they are exactly the same). In other words, each data is similar only unto itself. Obviously the *identity relation* [Intan, Mukaidono, 2000d], [Shenoi, Melton, 1989] is used to represent relationships between data. Consequently, two objects are considered as redundant objects if they exactly have the same data for all domain attributes. However, when we consider providing a more realistic view in representing data, we have to realize that data are often imprecise. Two distinct data in a given domain attribute might have degree of similarity between 0 and 1. An object is considered as an α-redundant object if there is another object to which the object is similar characterized by some domain attributes at least with the degree of α as shown in Definition 3.7.2. Moreover, we think of α as a set of the degree from 0 to 1 which corresponds to the set of domain attributes. This gives flexibility in determining the degree of α depending on type of data in domain attributes.

Example 3.7.1 Simply, Reproduction of Animals in Table 3.3 is recalled with corresponding to information system $I(U, A)$, where
$U = \{u_1, u_2, u_3\}$,
$A = \{Description(D), Reproduction(R)\}$.

Let us suppose for simplifying the problem degree of $\alpha = \{1, 1\}$, where $\alpha_D = 1$ and $\alpha_R = 1$. First, we determine similarity of data, for instance relationship between Horse and Mammals is discussed as follows.

$R(D(u_1), D(u_2)) = R(Horse, Mammals) = \text{P}(Mammals \to Horse) << 1$ or by conditional propositions p [Klir, Yuan, 1995], the relation is expressed by

p : If *Mammals* (*is true*) then *Horse* (*is true*).

It is not exactly true, because *Mammals* is not a part of *Horse*. On the other hand, $R(D(u_2), D(u_1)) = R(Mammals, Horse) = \text{P}(Horse \to Mammals) = 1$ or by conditional propositions p, the relation is expressed by

p : If *Horse* (*is true*) then *Mammals* (*is true*).

It is exactly true, because *Horse* is a part of *Mammals*. Next similarity classes of all objects are constructed by (3.39) after determined similarity of data for value of all domain attributes corresponding to the objects.

$S_A^\alpha(u_1) = \{u_1\}$,

$S_A^\alpha(u_2) = \{u_1, u_2\}$,

$S_A^\alpha(u_3) = \{u_3\}$.

It is obviously seen that $S_A^\alpha(u_1) \subset S_A^\alpha(u_2)$. By (3.40), we conclude that u_1 is a redundant object because it is covered by u_2.

3.8 Dependency of Domain attributes

Dependency of domain attributes is one of the important issues in the application of KDD and design of database such as recognizing partial and total dependencies as well as functional dependencies in relational database, determining redundant domain attributes in decision table, and others. Intuitively, a set of domain attributes D depends totally on C if all value of domain attributes from D are uniquely determined by values of domain attributes from C [Komorowski, Pawlak, Polkowski, Skowron, 1999]. In this section, we propose a concept for determining dependency of domain attributes based on α-coverings of the universe constructed by similarity classes as defined in Definition 3.7.1. Formally, the concept of dependency domain attributes can be defined in the following definition.

Definition 3.8.1 Let $I = (U, A)$ be an information system, where $U = \{u_1, ..., u_n\}$. For $C, D \subseteq A$; $S_C^{\beta}(u_1)$ and $S_D^{\gamma}(u_1)$ are defined as the sets that are similar to u_i by employing C to degrees of at least β and by employing D to degrees of at least γ, respectively in which β corresponds to C and γ corresponds to D. $\delta_i^{\beta,\gamma}(C,D)$ is defined as degrees of dependency C determines D, in object u_i by:

$$\delta_i^{\beta,\gamma}(C,D) = \frac{|S_C^{\beta}(u_i) \cap S_D^{\gamma}(u_i)|}{|S_C^{\beta}(u_i)|}. \tag{3.41}$$

From the universal set of objects U, we have a family of values $\{\delta_i^{\beta,\gamma}(C,D) \mid \forall i \in \mathrm{N}_n\}$. In general, by using this family of values, degree of dependency, C determines D, may be defined into three definitions as the following:

$$(minimum) \quad \delta_{\min}^{\beta,\gamma}(C,D) = \min_{i \in \mathrm{N}_n} \delta_i^{\beta,\gamma}(C,D), \tag{3.42}$$

$$(maximum) \quad \delta_{\max}^{\beta,\gamma}(C,D) = \max_{i \in \mathrm{N}_n} \delta_i^{\beta,\gamma}(C,D), \tag{3.43}$$

$$(average) \quad \delta_{avg}^{\beta,\gamma}(C,D) = \underset{i \in \mathrm{N}_n}{avg}\, \delta_i^{\beta,\gamma}(C,D), \tag{3.44}$$

By definition we can obtain some properties such as: if *D* depends totally on *C* then $R_C^{\beta}(u_i) \subseteq R_D^{\gamma}(u_i)$ for all *i* or all objects in *U*. It also means that the similarity classes generated by *C* are finer than the similarity classes generated by *D*. Also there are some properties such as:

- $\delta_{\min}^{\beta,\gamma}(C,D)=1 \Leftrightarrow \delta_{avg}^{\beta,\gamma}(C,D)=1$, and similarly
  $\delta_{\min}^{\beta,\gamma}(C,D)<1 \Leftrightarrow \delta_{avg}^{\beta,\gamma}(C,D)<1.$
- If $\delta_{\min}^{\beta,\gamma}(C,D)=1 \Leftrightarrow \delta_{avg}^{\beta,\gamma}(C,D)=1$, then we say that *D* depends totally on *C*.
  Otherwise, if $\delta_{avg}^{\beta,\gamma}(C,D)<1$, we say that *D* depends partially (in a degree $\delta_{avg}^{\beta,\gamma}(C,D)$) on *C*.
- Likewise, if $\delta_{i}^{\beta,\gamma}(C,D)=1$ we say that *D* depends totally on *C* in object u_i.
  Otherwise if $\delta_{i}^{\beta,\gamma}(C,D)<1$, we say that *D* depends partially (in a degree $\delta_{i}^{\beta,\gamma}(C,D)$) on *C* in object u_i.

Moreover, the concept of dependency discussed above corresponds to the concept of fuzzy functional dependency in the presence of fuzzy relational database [Intan, Mukaidono, 2000a], [Intan, Mukaidono, 2000d], [Intan, Mukaidono, 2004]. We may say that fuzzy functional dependency (FFD) *C* determines *D*, denoted by $C \rightarrow D$, holds in a relation *r* (or in the information system I) iff:

$$\delta_{i}^{\beta,\gamma}(C,D) \geq \delta_{i}^{\beta,\gamma}(D,C), \quad \text{for all } i \in \mathrm{N}_n. \tag{3.45}$$

It can be proved that the FFD satisfies Armstrong's Axioms [Armstrong, 1974], such that for *B, C, D* ∈ *A,* where *A*: set of domain attributes,

1. Reflexivity: $D \subseteq C \Rightarrow C \rightarrow D$,
2. Augmentation: $(C \rightarrow D, B \subseteq A) \Rightarrow CB \rightarrow D$,
3. Transitivity: $(C \rightarrow D \text{ and } D \rightarrow B) \Rightarrow C \rightarrow B$.

There are redundant domain attributes in decision table if there is a set of $C' \subset C$ such that:

$$\delta_i^{\beta',\gamma}(C',D) \geq \delta_i^{\beta,\gamma}(C,D), \quad \text{for all } i \in \mathrm{N}_n, \tag{3.46}$$

where set of $C - C'$ is considered as set of redundant domain attributes.

Example 3.8.1 Simply, we give an example of crisp data by considering an information system $I(U, A)$ as shown in Table 3.4, where $U = \{u_1, ..., u_8\}$ is a set of objects and $A = \{c_1, c_2, c_3, d_1\}$ is a set of domain attributes.

Table 3.4 $I(U, A = \{c_1, c_2, c_3, d_1\})$

U	c_1	c_2	c_3	d_1
u_1	w_1	x_1	y_1	z_1
u_2	w_1	x_2	y_3	z_1
u_3	w_2	x_2	y_3	z_2
u_4	w_1	x_1	y_1	z_2
u_5	w_2	x_2	y_1	z_2
u_6	w_1	x_1	y_1	z_1
u_7	w_2	x_2	y_3	z_1
u_8	w_1	x_2	y_3	z_1

Let us suppose that we would like to know the degree of dependency: C determines D, where $C = \{c_1, c_2, c_3\}$ and $D = \{d_1\}$ such that $C, D \subset A$. First, we calculate $\delta_1^{\beta,\gamma}(C,D)$ which means degree of dependency, C determines D in object u_1, where let us suppose (in crisp data) $\beta = \{1, 1, 1\}$ and $\gamma = \{1\}$ as follows.

$$\left.\begin{array}{l} S_C^{\beta}(u_i) = \{u_1, u_4, u_6\} \\ S_D^{\gamma}(u_i) = \{u_1, u_2, u_6, u_7, u_8\} \end{array}\right\} \Rightarrow \delta_1^{\beta\gamma}(C,D) = \frac{|\{u_1, u_6\}|}{|\{u_1, u_4, u_6\}|} = \frac{2}{3},$$

By using the same way, we calculate the degree of dependency that C determines D, for other objects: $\delta_2^{\beta\gamma}(C,D) = 1$, $\delta_3^{\beta\gamma}(C,D) = 1/2$,

$\delta_4^{\beta\gamma}(C,D)=1/3$, $\delta_5^{\beta\gamma}(C,D)=1$, $\delta_6^{\beta\gamma}(C,D)=2/3$, $\delta_7^{\beta\gamma}(C,D)=1/2$ and $\delta_8^{\beta\gamma}(C,D)=1$. Finally by (3.42), (3.43) and (3.44), we calculate minimum, maximum and average degree of dependency as given by:

$$\delta_{\min}^{\beta\gamma}(C,D)=1/2,\quad \delta_{\max}^{\beta\gamma}(C,D)=1,\quad \delta_{avg}^{\beta\gamma}(C,D)=17/24.$$

Conversely, we calculate degree of dependency, D determines C, for all objects as follows: $\delta_1^{\beta\gamma}(D,C)=2/5$, $\delta_2^{\beta\gamma}(D,C)=2/5$, $\delta_3^{\beta\gamma}(D,C)=1/3$, $\delta_4^{\beta\gamma}(D,C)=1/3$, $\delta_5^{\beta\gamma}(D,C)=1/3$, $\delta_6^{\beta\gamma}(D,C)=2/5$, $\delta_7^{\beta\gamma}(D,C)=1/5$ and $\delta_8^{\beta\gamma}(D,C)=2/5$. If we compare degree of dependency, *C* determines *D*, with degree of dependency, *D* determines *C*, it will be clearly seen that $\delta_i^{\beta\gamma}(C,D)\geq\delta_i^{\beta\gamma}(D,C)$ for all *i* or all objects, where it satisfies (3.45). Finally, we conclude that $C\rightarrow D$ holds in Table 3.4.

Again, let us suppose $C'=\{c_1, c_3\}$ in which $C'\subset C$. By (3.41), we calculate the degree of dependency, *C* determines *D*, for all objects: $\delta_1^{\beta\gamma}(C',D)=2/3$, $\delta_2^{\beta\gamma}(C',D)=1$, $\delta_3^{\beta\gamma}(C',D)=1/2$, $\delta_4^{\beta\gamma}(C',D)=1/3$, $\delta_5^{\beta\gamma}(C',D)=1$, $\delta_6^{\beta\gamma}(C',D)=2/3$, $\delta_7^{\beta\gamma}(C',D)=1/2$ and $\delta_8^{\beta\gamma}(C',D)=1$. These give exactly the same results as degree of dependency *C* determines *D*. By (3.46), $C-C'=\{c_2\}$ is a redundant domain attribute.

3.9 Conclusion

Weak fuzzy similarity relation was proposed as a generalization of fuzzy similarity relation [Zadeh, 1970]. Conditional probability relation was regarded as a concrete example of the weak fuzzy similarity relation. A generalization of classical rough set was proposed based on covering of the universe induced by conditional probability relation. Considering the rough sets approximation, two interpretations of lower and upper approximation operators are introduced in the presence of α-coverings, where α indicates the degree of similarity relationships between elements or objects in order to construct the covering. Some properties related to the lower and upper approximation operators are also examined. Rough membership functions are re-defined into three values: minimum, maximum and average. Four regions of a given set were provided and defined. Some properties were proposed related to set-theoretic operators. Also, two important applications were discussed in the relation to information system. A concept of determining α-redundant objects was introduced based on the concept of α-coverings of the universe. The concept of determining α-redundant objects is very important in order to reduce the number of decision rules in decision table. Finally, a concept of dependency of domain attributes was also proposed. Dependency of domain attributes is one of the important issues in the application of KDD and design of database such as recognizing partial and total dependencies, determining redundant domain attributes in the information system, and others. In addition, fuzzy functional dependency (FFD) as an important method in analyzing fuzzy relational database was defined based on dependency of domain attributes.

Chapter 4

Multi Rough Sets based on Multi-Context of Attributes

4.1 Introduction

In the real application, depending on the context, a given object may have different values of attributes. In other words, we may represent set of attributes based on different context, where they may provide different values for a given object. Context can be viewed as background or situation in which somehow we need to group some attributes as a subset of attributes and consider the subset as a context. For example, let us consider humans as a universal set of objects. Every person (object) might be characterized by some sets of attributes corresponding to some contexts such as his or her status as student, employeee, family member, club member, etc. In the context of student, his or her set of attributes might be {*ID-Number*, *Name*, *Address*, *Supervisor*, *Major*, etc.}. We may consider different sets of attributes in the relation to the contexts of both employee and family member. Still using example of humans as objects, especially for fuzzy data or perception-based data, set of attributes such as *height*, *weight* and *age*, might have different values for a given object depending on viewpoints (contexts) of American, Japanese and so on. For instance, Japanese may consider height of 175 cm as {*high*}, but American may consider it as {*medium*}. Therefore, it is necessary to consider multi-contexts information system as an extension of information system (see Section 4.2). Related to the rough set, every context as a set of attributes provides a partition of objects. Consequently, since n contexts (n subsets of attributes) provide n partitions, a given set of object, X, may then be represented into n pairs of lower and upper approximations defined as *multi-rough sets* of X. Related to the multi-rough sets, some properties and operations are proposed and examined. Primary concern is given to the generalization of contexts in the presence of multi-contexts information system. Three kinds of general contexts, namely AND-general context, OR-general context and OR^+ general context, are proposed. We show that AND-general context and OR^+-general context provide (disjoint) partitions but OR^+-general context provides covering of the universe. Then, a summarized rough set of a given crisp set of objects is able to be derived from partitions as well as covering of the general contexts. Finally, relations among three general contexts are examined and summarized.

4.2 Multi-Contexts Information System

A *Multi Rough Sets* is proposed based on multi-contexts of attributes, where every context is considered as a set of attributes. Partitions of multi-contexts are generated from a multi-contexts information system. Formally, the multi-contexts information system is defined by a pair $I = (U, \mathbf{A})$, where U is a universal set of objects and $\mathbf{A}$ is a non-empty set of contexts such as $\mathbf{A} = \{A_1, ..., A_n\}$. $A_i \in \mathbf{A}$ is a set of attributes and denoted as a context. Every attribute, $a \in A_i$, is associated with a set of V_a as its values called domain of a. It is NOT necessary for $i \neq j \Rightarrow A_i \cap A_j = \emptyset$. Attributes such as *height* and *weight* might belong to different contexts (i.e. American and Japanese) in which they may provide different values of certain attribute concerning a given object. Therefore, for $x \in U$, $a(x)^i \in V_a$ is denoted as the value of attribute a for objects x in the context $a \in A_i$. An indiscernibility relation (equivalence relation) is then defined in terms of context A_i such as for $x, y \in U$:

$$R_{A_i}(x,y) \Leftrightarrow a(x)^i = a(y)^i, \quad a(x)^i, a(y)^i \in V_a, \quad \forall a \in A_i, \quad (4.1)$$

Equivalence class of $x \in U$ in the context A_i is given by

$$[x]_{A_i} = \{y \in U \mid R_{A_i}(x,y)\}. \quad (4.2)$$

It should be verified that for $i \neq j$, $\exists x \in U$, $[x]_{Ai} \neq [x]_{Aj}$, otherwise A_i and A_j are redundant in term of providing similar partitions. By eliminating all redundant contexts, the number of contexts in the relation to the number of objects are satisfied the following equation.

$$\text{For} \mid U \mid = m, \ \mid \mathbf{A} \mid \leq B(m), \ \ B(m) = \sum_{i=0}^{m-1} C(m-1, i) \times B(i), \quad (4.3)$$

where $B(0) = 1$ and $C(n, k)$ is combination of size k from n elements given by:

$$C(n,k) = \frac{n!}{k!(n-k)!}.$$

$|U|$ and |A| are cardinalities of U and **A** representing the number of objects and the number of contexts, respectively. From set of contexts **A**, set of partitions of universal objects are derived and given by $\{U/A_1, ..., U/A_n\}$, where U/A_i as a partition of the universe based on context A_i contains all equivalence classes of $[x]_{Ai}$, $x \in U$.

4.3 Multi-Rough Sets

A multi-rough sets is defined as an approximate representation of a given crisp set of objects in the presence of a set of partitions derived from multi-context information systems providing set of rough sets corresponding to the set of partitions. Here, the multi-rough sets may be provided regardless of redundant contexts in multi-contexts information system. Clearly, every element of the multi-rough sets is a pair of lower and upper approximation corresponding to a given context. Formally, definition of multi-rough sets is defined as follows.

Definition 4.3.1 Let U be a non-empty universal set of objects. R_{A_i} and U/R_{A_i} are equivalence relation and partition with respect to set of attributes in the context of A_i. For $X \subseteq U$, corresponding to a set of contexts, $\mathbf{A}=\{A_1, A_2, \ldots, A_n\}$, $\mathbf{X}$ is defined as multi-rough sets of X as given as follows.

$$\mathbf{X} = \{(Lo(X_1), Up(X_1)), (Lo(X_2), Up(X_2)), \ldots, (Lo(X_n), Up(X_n))\}. \quad (4.4)$$

Thus, an element $(Lo(X_i), Up(X_i))$ of multi-rough sets is a pair of sets, lower and upper approximations in terms of context A_i. Similar to the definition of rough set, $Lo(X_i)$ and $Up(X_i)$ are defined by

$$Lo(X_i) = \{u \in U \mid [u]_{A_i} \subseteq X\} = \bigcup \{[u]_{A_i} \in U / A_i \mid [u]_{A_i} \subseteq X\}, \quad (4.5)$$

$$Up(X_i) = \{u \in U \mid [u]_{A_i} \cap X \neq \emptyset\} = \bigcup \{[u]_{A_i} \in U / A_i \mid [u]_{A_i} \cap X \neq \emptyset\}, \quad (4.6)$$

respectively. Similar to *bags* (multi-set) as proposed in [Yager, 1990], a multi-rough sets, $\mathbf{X}$, is characterized by a counting function $\Sigma_{\mathbf{X}}$ such that:

$$\sum\nolimits_{\mathbf{X}} \mathcal{P}(U)^2 \rightarrow \mathrm{N}, \quad (4.7)$$

where N is a set of non-negative integers and P(U) is power set of U. Basically, for any pair of lower and upper approximations $(M,N) \in \mathrm{P(U)}^2$, $\Sigma_{\mathbf{X}}((M,N))$ counts number of occurrences the pair (M,N) in the multi-rough sets $\mathbf{X}$, where it should be clarified that

$$(M, N) \notin \mathbf{X} \Rightarrow \sum\nolimits_{\mathbf{X}}((M, N)) = 0.$$

Also, a support set of $\mathbf{X}$ denoted by $\mathbf{X}^*$ is defined by satisfying the following equation:

$$(M, N) \in \mathbf{X}^* \Leftrightarrow \sum\nolimits_{\mathbf{X}}((M, N)) > 0, \quad (4.8)$$

where $\forall (M, N) \in \mathbf{X}^*, \sum\nolimits_{\mathbf{X}^*}((M, N)) = 1$. It can be proved that if $\mathbf{X}=\mathbf{X}^*$ then set of contexts $\mathbf{A}$ is free from redundancy, not vice versa. Some basic relations and operations are defined concerning sets of pair lower and upper approximations as elements of multi-rough sets. For $\mathbf{X}$ and $\mathbf{Y}$ are two multi-rough sets on U drawn from multi-contexts information system $\mathbf{A}$, where $|\mathbf{A}|=n$:

i. Containment

$\mathbf{X} \subseteq \mathbf{Y} \Leftrightarrow (Lo(X_i) \subseteq Lo(Y_i), Up(X_i) \subseteq Up(Y_i)), \forall i \in \mathrm{N}_n;$

ii. Equality: $\mathbf{X} = \mathbf{Y} \Leftrightarrow (Lo(X_i) = Lo(Y_i), Up(X_i) = Up(Y_i)), \forall i \in \mathrm{N}_n;$

iii. Complement:

$\mathbf{Y} = \neg \mathbf{X} \Leftrightarrow (Lo(Y_i) = U - Up(X_i), Up(Y_i) = U - Lo(X_i)), \forall i \in \mathrm{N}_n;$

iv. Union: $\mathbf{X} \cup \mathbf{Y} \Leftrightarrow \{(Lo(X_i) \cup Lo(Y_i), Up(X_i) \cup Up(Y_i)) \mid \forall i \in \mathrm{N}_n\};$

v. Intersection:

$\mathbf{X} \cap \mathbf{Y} \Leftrightarrow \{(Lo(X_i) \cap Lo(Y_i), Up(X_i) \cap Up(Y_i)) \mid \forall i \in \mathrm{N}_n\};$

where N_n is a set of non-negative integers less or equal to n. Obviously, the operations given in (i)-(v) are strongly related to the order of elements corresponding to set of contexts. Related to the occurrence of elements and despite the order of elements in multi-rough sets, we may consider the following basic operations.

a) Containment: $\mathbf{X} \prec \mathbf{Y} \Leftrightarrow \sum_{\mathbf{X}}((M,N)) \leq \sum_{\mathbf{Y}}((M,N)), \forall (M,N);$

b) Equality: $\mathbf{X} \equiv \mathbf{Y} \Leftrightarrow \sum_{\mathbf{X}}((M,N)) = \sum_{\mathbf{Y}}((M,N)), \forall (M,N);$

c) Union: $\sum_{\mathbf{X} \oplus \mathbf{Y}}((M,N)) = \mathbf{max}[\Sigma_{\mathbf{X}}((M,N)), \Sigma_{\mathbf{Y}}((M,N))], \forall (M,N);$

d) Intersection:

$\sum_{\mathbf{X} \otimes \mathbf{Y}}((M,N)) = \mathbf{min}[\Sigma_{\mathbf{X}}((M,N)), \Sigma_{\mathbf{Y}}((M,N))], \forall (M,N);$

e) Insertion:

$\sum_{\mathbf{X} + \mathbf{Y}}((M,N)) = \Sigma_{\mathbf{X}}((M,N)) + \Sigma_{\mathbf{Y}}((M,N)), \forall (M,N) \in \mathbf{X} \oplus \mathbf{Y};$

f) Minus:

$\sum_{\mathbf{X} - \mathbf{Y}}((M,N)) = \mathbf{max}[\Sigma_{\mathbf{X}}((M,N)) - \Sigma_{\mathbf{Y}}((M,N)), 0], \forall (M,N) \in \mathbf{X};$

The above basic operations satisfy some properties as the following:

1. Idempotent laws:

   $\mathbf{X} \cup \mathbf{X} = \mathbf{X}, \quad \mathbf{X} \cap \mathbf{X} = \mathbf{X}, \quad \mathbf{X} \oplus \mathbf{X} = \mathbf{X}, \quad \mathbf{X} \otimes \mathbf{X} = \mathbf{X};$

2. Commutative laws:

   $\mathbf{X} \cup \mathbf{Y} = \mathbf{Y} \cup \mathbf{X}, \quad \mathbf{X} \cap \mathbf{Y} = \mathbf{X} \cap \mathbf{Y}, \quad \mathbf{X} \oplus \mathbf{Y} = \mathbf{Y} \oplus \mathbf{X},$

   $\mathbf{X} \otimes \mathbf{Y} = \mathbf{Y} \otimes \mathbf{X}, \quad \mathbf{X} + \mathbf{Y} = \mathbf{X} + \mathbf{Y};$

3. Associative laws:

   $\mathbf{W} \cup (\mathbf{X} \cup \mathbf{Y}) = (\mathbf{W} \cup \mathbf{Y}) \cup \mathbf{X}, \quad \mathbf{W} \cap (\mathbf{X} \cap \mathbf{Y}) = (\mathbf{W} \cap \mathbf{X}) \cap \mathbf{Y},$

$$\mathbf{W}\oplus(\mathbf{X}\oplus\mathbf{Y})=(\mathbf{W}\oplus\mathbf{Y})\oplus\mathbf{X}, \quad \mathbf{W}\otimes(\mathbf{X}\otimes\mathbf{Y})=(\mathbf{W}\otimes\mathbf{Y})\otimes\mathbf{X},$$
$$\mathbf{W}+(\mathbf{X}+\mathbf{Y})=(\mathbf{W}+\mathbf{Y})+\mathbf{X};$$

4. Absorption laws:
$$\mathbf{X}\cap(\mathbf{X}\cup\mathbf{Y})=\mathbf{X}, \quad \mathbf{X}\cup(\mathbf{X}\cap\mathbf{Y})=\mathbf{X},$$
$$\mathbf{X}\otimes(\mathbf{X}\oplus\mathbf{Y})=\mathbf{X}, \quad \mathbf{X}\oplus(\mathbf{X}\otimes\mathbf{Y})=\mathbf{X},$$
$$\mathbf{X}\otimes(\mathbf{X}+\mathbf{Y})=\mathbf{X}, \quad \mathbf{X}\oplus(\mathbf{X}+\mathbf{Y})=\mathbf{X}+\mathbf{Y};$$
5. Distributive laws:
$$\mathbf{W}\cup(\mathbf{X}\cap\mathbf{Y})=(\mathbf{W}\cup\mathbf{X})\cap(\mathbf{W}\cup\mathbf{Y}),$$
$$\mathbf{W}\cap(\mathbf{X}\cup\mathbf{Y})=(\mathbf{W}\cap\mathbf{X})\cup(\mathbf{W}\cap\mathbf{Y}),$$
$$\mathbf{W}\oplus(\mathbf{X}\otimes\mathbf{Y})=(\mathbf{W}\oplus\mathbf{X})\otimes(\mathbf{W}\oplus\mathbf{Y}),$$
$$\mathbf{W}\otimes(\mathbf{X}\oplus\mathbf{Y})=(\mathbf{W}\otimes\mathbf{X})\oplus(\mathbf{W}\otimes\mathbf{Y}),$$
$$\mathbf{W}+(\mathbf{X}\otimes\mathbf{Y})=(\mathbf{W}+\mathbf{X})\otimes(\mathbf{W}+\mathbf{Y}),$$
$$\mathbf{W}+(\mathbf{X}\oplus\mathbf{Y})=(\mathbf{W}+\mathbf{X})\oplus(\mathbf{W}+\mathbf{Y});$$
6. Additive laws:
$$\mathbf{X}\oplus\mathbf{Y}=\mathbf{X}+\mathbf{Y}-(\mathbf{X}\otimes\mathbf{Y});$$
7. Double negation law:
$$\neg\neg\mathbf{X}=\mathbf{X};$$
8. De Morgan laws:
$$\neg(\mathbf{X}\cap\mathbf{Y})=\neg\mathbf{X}\cup\neg\mathbf{Y}, \quad \neg(\mathbf{X}\cup\mathbf{Y})=\neg\mathbf{X}\cap\neg\mathbf{Y};$$
9. Maximum multi-rough sets (**U**={(*U*,*U*),…}, |**U**|=|**X**|);
$$\mathbf{X}\cap\mathbf{U}=\mathbf{X}, \quad \mathbf{X}\cup\mathbf{U}=\mathbf{U};$$
10. Minimum multi-rough sets (**E**={(∅,∅),…}, |**E**|=|**X**|);
$$\mathbf{X}\cap\mathbf{E}=\mathbf{E}, \quad \mathbf{X}\cup\mathbf{E}=\mathbf{X};$$
11. Kleene's laws:
$$(\mathbf{X}\cap\neg\mathbf{X})\cap(\mathbf{Y}\cup\neg\mathbf{Y})=(\mathbf{X}\cap\neg\mathbf{X}), \quad (\mathbf{X}\cap\neg\mathbf{X})\cup(\mathbf{Y}\cup\neg\mathbf{Y})=(\mathbf{Y}\cup\neg\mathbf{Y});$$

Since basic operations defined in (iii)-(v) do not satisfy complementary laws ($\mathbf{X}\cap\neg\mathbf{X})\neq\mathbf{E}$ and $\mathbf{X}\cup\neg\mathbf{X})\neq\mathbf{U}$), they do not satisfy Boolean algebra but just Kleene algebra instead. When union and intersection are applied for all pair elements of multi-rough sets **X**, we have:

$$\Gamma(X) = \bigcup_i Up(X_i), \tag{4.9}$$

$$\Theta(X) = \bigcap_i Up(X_i), \tag{4.10}$$

$$\Phi(X) = \bigcup_i Lo(X_i), \tag{4.11}$$

$$\Psi(X) = \bigcap_i Lo(X_i), \tag{4.12}$$

where $Up(\mathbf{X}) = \{(\Gamma(X), \Theta(X))\}$ and $Lo(\mathbf{X}) = \{(\Phi(X), \Psi(X))\}$ are defined as summary multi-rough sets in which they have only one pair element. Their relationship can be easily verified by:

$$\Psi(X) \subseteq \Phi(X) \subseteq X \subseteq \Theta(X) \subseteq \Gamma(X),$$

where we may consider pair of $(\Phi(X),\ \Theta(X))$ as a finer approximation and pair of $(\Psi(X),\ \Gamma(X))$ as a worse approximation of $X \subseteq U$. From the definition of summary multi-rough sets, it satisfies some properties such as:

(1) $X \subseteq Y \Leftrightarrow [\Psi(X) \subseteq \Psi(Y), \Phi(X) \subseteq \Phi(Y), \Theta(X) \subseteq \Theta(Y), \Gamma(X) \subseteq \Gamma(Y)]$,

(2) $\Psi(X) = \neg\Gamma(\neg X), \Phi(X) = \neg\Theta(\neg X), \Theta(X) = \neg\Phi(\neg X), \Gamma(X) = \neg\Psi(\neg X)$,

(3) $\Psi(U) = \Phi(U) = \Theta(U) = \Gamma(U) = U, \Psi(\emptyset) = \Phi(\emptyset) = \Theta(\emptyset) = \Gamma(\emptyset) = \emptyset$,

(4) $\Psi(X \cap Y) = \Psi(X) \cap \Psi(Y), \Phi(X \cap Y) = \Phi(X) \cap \Phi(Y), \Theta(X \cap Y) \leq \Theta(X) \cap \Theta(Y), \Gamma(X \cap Y) \leq \Gamma(X) \leq \Gamma(Y)$,

(5) $\Psi(X \cup Y) \geq \Psi(X) + \Psi(Y) - \Psi(X \cap Y), \Phi(X \cup Y) \geq \Phi(X) + \Phi(Y) - \Phi(X \cap Y), \Phi(X \cup Y) \leq \Phi(X) + \Phi(Y) - \Phi(X \cap Y), \Gamma(X \cup Y) \leq \Gamma(X) + \Gamma(Y) - \Gamma(X \cap Y)$,

Special consideration is given to the following two characteristics of context.

1. A_i is called *total ignorance* (τ) if $x \in U$, $[x]_\tau = U$.
   Therefore $\forall X \subseteq U, X \neq \emptyset \Rightarrow Lo(X_\tau) = \emptyset, Up(X_\tau) = U$.
2. A_i is called *identity* (ι) if $\forall x \in U$, $[x]_\iota = \{x\}$.
   Therefore, $\forall X \subseteq U \Rightarrow Lo(X_\iota) = Up(X_\iota) = X$.

Obviously, related to union and intersection operations, we have the following properties:

$\forall A_i \in \mathbf{A}, X \subseteq U,$

- Union: $X \neq \emptyset \Rightarrow Up(X_i) \cup Up(X_\tau) = U,\ Lo(X_i) \cup Lo(X_\tau) = Lo(X_i),$
  $Up(X_i) \cup Up(X_\iota) = Up(X_i),\ \ Lo(X_i) \cup Lo(X_\iota) = X.$
- Intersection: $Up(X_i) \cap Up(X_\tau) = Up(X_i),\ Lo(X_i) \cap Lo(X_\tau) = \emptyset,$
  $Up(X_i) \cap Up(X_\iota) = X_i,\ \ Lo(X_i) \cap Lo(X_\iota) = Lo(X_i).$

From the relation with union and intersection operations, τ is the identity context for union operation of lower approximation as well as for intersection operation of upper approximation. On the other hand, ι is the identity context for union operation of upper approximation as well as for intersection operation of lower approximation.

Furthermore, in order to characterize multi-rough sets based on the number of objects (elements of U), two count functions are defined as follows:

Definition 4.3.2 $\eta_{\mathbf{X}} : U \rightarrow \mathrm{N}_n$ and $\sigma_{\mathbf{X}} : U \rightarrow \mathrm{N}_n$ are defined as two functions to characterize multi-rough set by counting total number of copies of a given element of U in upper and lower sides of multi rough set **X**, respectively, as given by:

$$\eta_{\mathbf{X}}(x) = \sum_{i}^{n} \theta_{Up(X_i)}(x), \tag{4.13}$$

$$\sigma_{\mathbf{X}}(x) = \sum_{i}^{n} \theta_{Lo(X_i)}(x), \tag{4.14}$$

where $|\mathbf{A}|=n$ and $\theta_M(x) = 1 \Leftrightarrow x \in M$, otherwise $\theta_M(x) = 0$.

These count functions are similar to one proposed in [Yager, 1990] talking about *bags* (multi-set). Similar results will be found by firstly taking *insertion operation* to all lower side yielding a multi-set of lower side as well as all upper side yielding a multi-set of upper side. Then, the counting function is used to calculate number of copies of each element in both multi-sets. Related to summary rough sets, these two count functions, η and σ, provide some properties such as for $X, Y \in U$, $|\mathbf{A}| = n$:

1. $\eta_{\mathbf{X}}(y) \geq \sigma_{\mathbf{X}}(y), \forall y \in U,$
2. $\sigma_{\mathbf{X}}(y) > 0 \Rightarrow y \in X,$
3. $y \in X \Rightarrow \eta_{\mathbf{X}}(y) = n,$
4. $y \in \Theta(X) \Leftrightarrow \eta_{\mathbf{X}}(y) = n,$
5. $y \in \Psi(X) \Leftrightarrow \sigma_{\mathbf{X}}(y) = n,$
6. $\eta_{\mathbf{X}}(y) > 0 \Leftrightarrow \Gamma(X),$
7. $\sigma_{\mathbf{X}}(y) > 0 \Leftrightarrow \Phi(X),$
8. $\mathbf{X} \subseteq \mathbf{Y} \Rightarrow \eta_{\mathbf{X}}(y) \leq \eta_{\mathbf{Y}}(y), \sigma_{\mathbf{X}}(y) \leq \sigma_{\mathbf{Y}}(y), \forall y \in U,$
9. $\mathbf{X} = \mathbf{Y} \Rightarrow \eta_{\mathbf{X}}(y) = \eta_{\mathbf{Y}}(y), \sigma_{\mathbf{X}}(y) = \sigma_{\mathbf{Y}}(y), \forall y \in U,$
10. $\eta_{\mathbf{X}\cup\mathbf{Y}}(y) = \eta_{\mathbf{X}}(y) + \eta_{\mathbf{X}}(y) - \eta_{\mathbf{X}\cap\mathbf{Y}}(y),$
11. $\sigma_{\mathbf{X}\cup\mathbf{Y}}(y) = \sigma_{\mathbf{X}}(y) + \sigma_{\mathbf{X}}(y) - \sigma_{\mathbf{X}\cap\mathbf{Y}}(y),$
12. $\eta_{\mathbf{X}\oplus\mathbf{Y}}(y) = \eta_{\mathbf{X}}(y) + \eta_{\mathbf{X}}(y) - \eta_{\mathbf{X}\otimes\mathbf{Y}}(y),$
13. $\sigma_{\mathbf{X}\oplus\mathbf{Y}}(y) = \sigma_{\mathbf{X}}(y) + \sigma_{\mathbf{X}}(y) - \sigma_{\mathbf{X}\otimes\mathbf{Y}}(y),$

Simply, by dividing the count functions with total number of contexts ($|\mathbf{A}| = n$), we define two membership functions, $\mu_{\mathbf{X}}(y) : U \rightarrow [0,1]$ and $v_{\mathbf{X}}(y) : U \rightarrow [0,1]$ by

$$\mu_{\mathbf{X}}(y) = \frac{\eta_{\mathbf{X}}(y)}{n}, \tag{4.15}$$

$$v_{\mathbf{X}}(y) = \frac{\sigma_{\mathbf{X}}(y)}{n}, \tag{4.16}$$

where $\mu_{\mathbf{X}}(y)$ and $v_{\mathbf{X}}(y)$ represent membership value of y in upper and lower multi-set $\mathbf{X}$, respectively. Actually, μ and ν are nothing but another representation of the count functions. However, we may consider pair of ($v_{\mathbf{X}}(y), \mu_{\mathbf{X}}(y)$) as an *interval membership function* of $y \in U$ in the presence of multi-contexts of attributes. Similarly, by changing n to 1 in Property number 3-5, μ and ν have exactly the same properties as given by η and σ, respectively.

4.4 Generalization of Contexts

Generalization of contexts means that all contexts of attributes are combined for the purpose of providing one general context. Here, we propose three kinds of general context, namely AND-general context, OR-general context and OR$^+$-general context.

First, general context provided by AND logic operator to all attributes of all contexts, called AND-*general context*, is simply constructed by collecting all elements of attributes of all contexts to the general context as defined by the following definition.

Definition 4.4.1 Let A = {A_1, A_2, . . . , A_n} be set of contexts. $A_\wedge$ is defined as AND-general context by $A_\wedge = A_1 \wedge A_2 \wedge \cdots \wedge A_n$, where $A_\wedge$ is a result of summation of all conditions as given by all attributes of A_i, $\forall i \in \mathrm{N}_n$ or simply,

$$A_\wedge = A_1 + A_2 + \cdots + A_n \qquad (4.17)$$

In Definition 4.4.1, nevertheless, it was defined before in Section 4.2 that it is not necessary $i \neq j \Rightarrow A_i \cap A_j = \emptyset$. Here, every attribute is regarded uniquely and independently in providing value of the attribute corresponding to a given object in terms of a certain context. It can be proved that $A_\wedge$ satisfies $|A_\wedge| = \sum_{i=1}^{n} |A_i|$. Also, $\forall [u]_{A_\wedge}, \forall i \in \mathrm{N}_n, \exists [u]_{A_i}$ such that $[u]_{A_\wedge} \subseteq [u]_{A_i}$. For a given $X \subseteq U$, $Lo(X_\wedge)$ and $Up(X_\wedge)$ are defined as lower and upper approximation of X provided by set of attributes, $A_\wedge$. Approximation space performed by AND-general context is regarded as the finest disjoined partition by combining all partition of contexts and considering every possible area of intersection among aquivalence classes as a equivalence class of AND-general context (see Figure 4.1.c). Therefore, it provides the finest approximation of rough set.

Second, if the relationships among contexts are operated by OR logic operator, the independency of every context persists in the process of generalization. Clearly, it provides a covering instead of a disjoint partition of the universal objects. Since, the general context providing covering [Intan, Mukaidono, Yao, 2001], [Yao, Zhang, 2000], it may also be called *Cover-general context* (*C-general context*, for short) and defined as follows.

Definition 4.4.2 Let A = {A_1, A_2, ..., A_n} be set of contexts. $A_\vee$ is defined as C-general context by: $A_\vee = A_1 \vee A_2 \vee \cdots \vee A_n$, such that

$$U / A_\vee = \bigcup_{i=1}^{n} U / A_i \tag{4.18}$$

where $U / A_\vee$ is a covering of the universe as union of all equivalence classes in $U / A_i, i \in \mathrm{N}_n$.

Consequently, $|U / A_\vee| \leq \sum_{i=1}^{n} |U / A_i|$ and $\forall C \in U / A_\vee, \forall i \in \mathrm{N}_n, \exists [u]_{A_i}$ such that $C = [u]_{A_i}$, where C is a *similarity class* in covering and $[u]_{A_i}$ is an equivalence class in the partition of U/A_i. We call C as a similarity class as a means to distinguish between equivalence class provided by equivalence relation as usually used in partition and one used in covering. Every similarity class might take overlap one to each other. A given object $u \in U$ possibly belongs to more than one similarity classes. It can be verified that for $X \subseteq U$, $Lo(X_\vee)$ and $Up(X_\vee)$, as a pair of lower and upper approximations of X in terms of $A_\vee$, can be defined by,

$$Lo(X_\vee) = \bigcup_{i=1}^{n} Lo(A_i), \tag{4.19}$$

$$Up(X_\vee) = \bigcup_{i=1}^{n} Up(A_i), \tag{4.20}$$

where $Lo(X_i)$ and $Up(X_i)$ are lower and upper approximation of X based on the context A_i. It can be proved that iterative operation is applied in the upper approximation operator as given by $Up(X_\vee) \subseteq Up(Up(X_\vee))$. We may then consider $M(Up(X_\vee))$ as a maximum upper approximation given by $Up(X_\vee) \subseteq Up(Up(X_\vee)) \subseteq \cdots \subseteq M(Up(X_\vee))$, where the iterative operation is no longer applied in the maximum upper approximation or $Up(M(Up(X_\vee))) = M(Up(X_\vee))$. Related to the covering of the univeral objects, some properties are given in [Intan, Yao, Mukaidono 2003a]. Moreover, related to the summary of multi-rough sets as defined in the previous section, we found that $Up(X_\vee) = \Gamma(X)$ and $Lo(X_\vee) = \Phi(X)$.

The third general context is called OR$^+$-*general context* in which transitive closure operation is applied to the covering as result of OR-general context or C-general context. In other words, equivalence classes of OR$^+$-general context are provided by union of all equivalence classes of all partitions (of all contexts) that overlap one to each other. Similarity classes of OR$^+$-general context is defined as the following definition.

Definition 4.4.3 Let A = {A_1, A_2, ..., A_n} be set of contexts. $A_\vee^+$ is defined as OR$^+$-general context by: $A_\vee^+ = A_1 \circ A_2 \circ \cdots \circ A_n$, such that $y \in [x]_{A_\vee^+}$ iff

$$(\exists C_i \in U / A_\vee,\ \ x, y \in C_i) \quad \text{OR} \quad (\exists C_{i1}, C_{i2}, \ldots, C_{im} \in U / A_\vee, \; x \in C_{i1}, C_{ik} \cap C_{i(k+1)} \neq \emptyset, k = 1, \ldots, m-1, y \in C_{im}). \tag{4.21}$$

where $m \leq n$ and $[x]_{A_\vee^+}$ is an equivalence class containing x in terms of $A_\vee^+$. For $U / A_\vee$ be a set of equivalence classes provided by all contexts, equivalence classes generated by $A_\vee^+$ are able to be constructed by the following algorithm:

$S_i \in U / A_\vee^+, i \in \mathrm{N}$: Equivalence classes of OR$^+$-general context.

p = 0; SC = $U / A_\vee$ *U/A∨*;

while SC = ∅ *do* {

  $p = p + 1$; S_p = ∅;

  SC = SC − {M}; M is an element (similarity class) of SC.

  $S_p = M$;

  SS = SC;

  while SS ≠ ∅ *do* {

SS = SS − $\{M\}$; M is an element
(similarity class) of SS.
if $S_p \cap M \neq \emptyset$ *then* {
$S_P = S_P \cup M$;
SC = SC − $\{M\}$;
}
}
}

Finally, by algorithm in Definition 4.4.3, there will be *p* equivalence classes. Possibly, *p* might be equal to 1 in case all elements in $U/A_\vee$ transitively join each other. It can be proved that all equivalence classes in $U/A_\vee^+$ are disjoint. Also, $\forall S \in U/A_\vee^+$ such that $\forall i \in \mathrm{N}_n, \exists M \in U/A_i, M \subseteq S$. For a given $X \subseteq U$, $Lo(A_\vee^+)$ and $Up(A_\vee^+)$ are defined as lower and upper approximation of X provided by set of attributes, $A_\vee^+$. Approximation space performed by OR^+-general context is regarded as the worst disjoined partition. Therefore, it provides the worst approximation of rough set. Related to the maximum upper approximation based on C-general context, it can be verified that $\mathrm{apr}Up(A_\vee^+) = M(Up(A_\vee))$. Compare to the summary of multi-rough sets and approximation based on AND-general context, we have

$$Lo(A_\vee^+) \subseteq \Psi(X) \subseteq \Phi(X) \subseteq Lo(A_\wedge) \subseteq X \subseteq Up(A_\wedge) \subseteq \Theta(X) \subseteq \Gamma(X) \subseteq Up(A_\vee^+)$$

How generalization of contexts applied in the approximation of *X* might be illustrated by the following figure. It is given two different contexts, A_1 and A_2 and their approximation of *X* as shown in Figure 4.1 (*a*) and (*b*). Approximations of *X* based on AND, OR and OR^+-general context are given in Figure 4.1 (*c*), (*d*) and (*e*), respectively.

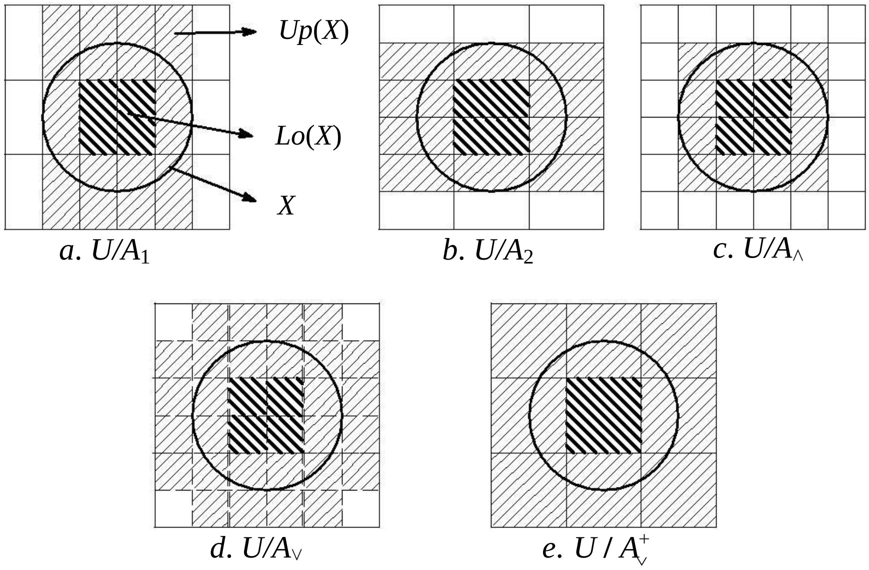

Figure 4.1 Generalization Contexts

4.5 Conclusion

This chapter proposed multi rough sets based on multi-context of attributes. Basic operations and some properties were examined. Two count functions as well as their properties ware defined and examined to characterize multi rough sets. Finally, we proposed three types of general contexts, namely AND-general context, C-general context and OR^{+}-general context. This chapter also discussed briefly relation among approximations provided by the general contexts. In the future work, we need to apply and implement the concept of multi rough sets in the real world application.

Chapter 5

Summary

In this book, we originally proposed and discussed some concepts related to rough sets, fuzzy sets and granular computing for the purpose of constructing intelligent information system as the following:

Chapter 2 discussed relationship between fuzziness and probability. We showed that fuzziness and probability played different roles in uncertainty. Therefore, they can be combined each other in order to represent probability of ill-defined event (called fuzzy event) in which the event can be represented by fuzzy set. Similar to fuzzy set, rough set regarded as another generalization of crisp set can be used to represent rough event in the relation to probability. We then examined their properties in the relation to belief and plausibility measures.

Chapter 3 gave a major contribution in generalization of rough sets induced by fuzzy conditional probability relation. Two applications, α-redundancy objects and dependency of domain attributes, were discussed in the presence of information system.

Still related to rough sets, Chapter 4 proposed a new concept of multi rough sets based on multi-contexts of attributes. Here, we need to find a real-world application to which the multi rough sets can be applied.

It should be mentioned that mostly the results of my research as discussed in this book is still in theory in which it is necessary to apply and implement them into the real-world applications. Also, in the future we need to consider the following topics of research in the relation to rough set.

- Completing generalization of rough sets induced by conditional probability relation.
- Rough Reasoning.
- Type-2 Rough Sets. Level *k* Rough Sets.
- Rough Measure.
- Rough Graph.

Bibliography

[1] Armstrong, W.W., ‘Dependency Structures of Database Relationship’, *Information Processing,* (1974), pp. 580-583.

[2] Dubois, D., Prade, H., *Fuzzy Sets and Systems: Theory and Applications*, (Academic Press, New York, 1980).

[3] Dubois, D., Prade, H., ’Rough Fuzzy Sets and Fuzzy Rough Set’, *International Journal of General Systems,* Vol. 17, (1990), pp. 191-209.

[4] Dubois, D., Prade, H., ‘Fuzzy Sets and Probability: Misunderstandings, Bridges and Gaps’, *Proc. Second IEEE Intern. Conf. on Fuzzy Systems, San Francisco*, (1993), pp. 1059-1068.

[5] Dempster, A.P., ’Upper and Lower Probability Induced by Multivalued Mappings’,

[6] *Annals of Mathematical Statistics, 38*, (1967), pp. 325-339.

[7] Intan, R., Mukaidono, M., ‘Application of Conditional Probability in Constructing Fuzzy Functional Dependency (FFD)’, *Proceedings of AFSS’00*, (2000a), pp. 271-276.

[8] Intan, R., Mukaidono, M., ‘A proposal of Fuzzy Functional Dependency based on Conditional Probability’, *Proceeding of FSS’00 (Fuzzy Systems Symposium),* (2000b), pp. 199-202.

[9] Intan, R., Mukaidono, M., ‘Fuzzy Functional Dependency and Its Application to Approximate Querying’, *Proceedings of IDEAS’00,* (2000c), pp. 47-54.

[10] Intan, R., Mukaidono, M., ‘Conditional Probability Relations in Fuzzy Relational Database’, *Proceedings of RSCTC’00, LNAI 2005,* Springer & Verlag, (2000d), pp. 251-260.

[11] Intan, R., Mukaidono, M, Yao, Y.Y., 'Generalization of Rough Sets with coverings of the Universe Induced by Conditional Probability Relations', *Proceedings of Inter-national Workshop on Rough Sets and Granular Computing, LNAI 2253,* Springer & Verlag, (2001a), pp. 311-315.
[12] Intan, R., Mukaidono, M, 'Dependency of Domain Attributes based on Fuzzy Conditional Probability Relations', *Proceeding of FSS'01 (Fuzzy Systems Symposium),* (2001b), pp. 563-566.
[13] Intan, R., Mukaidono, M, 'Generalized Rough Membership Function with α-Coverings of the Universe induced by Conditional Probability Relation', *Proceeding of FAN'01,* (2001c), pp. 557 560.
[14] Intan, R., Mukaidono, M., 'Redundant Object and Dependency of Domain Attributes in *α*-Coverings of the Universe', *Proceedings of FUZZ-IEEE,* (2001d), pp. 1444-1447.
[15] Intan, R., Mukaidono, M., 'Approximate Data Querying induced by Fuzzy Conditional Probability Relation', *Proceeding of Vietnam-Japan bilateral Symposium on Fuzzy Systems,* (2001e), pp. 140-147.
[16] Intan, R., Mukaidono, M., 'Degree of Similarity in Fuzzy Partition', *Proceedings of AFSS'02, LNAI 2275,* Springer & Verlag, (2002a), pp. 20-26.
[17] Intan, R., Mukaidono, M., 'Generalization of Rough Membership Function based on *α*-Coverings of the Universe', *Proceedings of AFSS'02, LNAI 2275,* Springer & Verlag, (2002b), pp.129-135.
[18] Intan, R., Mukaidono, M., Emoto, M., 'Knowledge Based Representation of Fuzzy Sets', *Proceeding of The 11th IEEE International Conference on Fuzzy Systems,* (2002c), pp. 590-595.
[19] Intan, R., Mukaidono, M., 'Approximate Reasoning in Knowledge-based Fuzzy Sets', *Proceeding of NAFIPS-FLINT 2002,* (2002d), IEEE Publisher, pp. 439-444.
[20] Intan, R., Mukaidono, M., 'A Proposal of Probability of Rough Event based on Probability of Fuzzy Event', *Proceedings of RSCTC'02, LNAI 2475,* Springer & Verlag, (2002e), pp. 357-364.

[21] Intan, R., Mukaidono, M., 'Generalization of Fuzzy Rough Sets by Fuzzy Covering Based On Weak Fuzzy Similarity Relation', *Proceeding of FSKD 2002,* (2002f), pp. 344-348.

[22] Intan, R., Mukaidono, M., 'Probability of Fuzzy Event to Probability of Rough Event', *Proceedings of FSKD 2002,* (2002g), pp. 549-553.

[23] Intan, R., Mukaidono, M., 'Generalization of Rough Sets and its Applications in Information System', *International Journal of Intelligent Data Analysis Vol. 6(4),* IOS Press, (2002h), pp. 323-339.

[24] Intan, R., Mukaidono, M., 'On Knowledge-based Fuzzy Sets', *International Journal of Fuzzy Systems, Vol. 4(2),* CFSAT, (2002i), pp. 655-664.

[25] Intan, R., Mukaidono, M., 'Approximate Data Querying in Fuzzy Relational Database', *Journal of Advanced Computational Intelligent Vol. 6(1),* (2002j), pp. 33-40.

[26] Intan, R., Mukaidono, M., 'Generalized Fuzzy Rough Sets By Conditional Probability Relations', *International Journal of Pattern Recognition and Artificial Intelligence Vol. 16(7),* World Scientific, (2002k), pp. 865-881.

[27] Intan, R., Yao, Y.Y., Mukaidono, M., 'Generalization of Rough Sets Using Weak Fuzzy Similarity Relations', *Rough Set Theory and Granular Computing, Advances in Soft Computing*, Physica-Verlag, (2003a), pp. 37-46.

[28] Intan, R., Mukaidono, M., 'Hybrid Probabilistic Models of Fuzzy and Rough Events', Journal of Advanced Computational Intelligent Vol. 7(3), (2003b), pp. 323-329.

[29] Intan, R., Mukaidono, M., 'From Evidence Theory to Probability of Generalized Fuzzy-Rough Events', *Proceedings of Fuzzy Information Processing,* (2003c), pp. 883-888.

[30] Intan, R., Mukaidono, M., 'Multi-Rough Sets Based on Multi-contexts of Attributes', *Proceedings of the 9th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing* (RSFDGrC), LNAI 2639, Springer-Verlag, (2003d), pp. 279-282.

[31] Intan, R., Mukaidono, M., ‘Multi-Rough Sets and Generalizations of Contexts in Multi-Contexts Information System’, *Proceedings of the 14th International Symposium on Methodologies for Intelligent Systems* (ISMIS), LNAI 2871, Springer-Verlag, (2003e), pp. 174-178.

[32] Intan, R., Mukaidono, M., ‘Fuzzy Conditional Probability Relations and its Applications in Fuzzy Information System’, *Knowledge and Information systems, an International Journal*, Springer-Verlag, (2004), pp. 345-365.

[33] Intan, R., Mukaidono, M., ‘On the Generalization of Fuzzy Rough Approximations Based on Asymmetric Relation’, *Computational Intelligence for Modelling and Prediction*, Netherland Springer-Verlag, (2005), pp. 73-88.

[34] Inuiguchi, M., Tanino, T., ‘On Rough Sets under Generalized Equivalence Relations’, *Proceedings of International Workshop on Rough Sets and Granular Computing*, (2001), pp. 167-171.

[35] Klir, G.J., Yuan, B., Fuzzy Sets and Fuzzy Logic: Theory and Applications, (Prentice Hall, New Jersey, 1995).

[36] Komorowski, J., Pawlak, Z., Polkowski, L., Skowron, A., ‘Rough Sets: A Tutorial’, *Rough Fuzzy Hybridization*, (S.K. Pal, A. Skowron, Eds.) (Springer, 1999), pp. 3-98.

[37] Kosko, B., ‘Fuzziness VS. Probability’, *International Journal of General Systems, Vol. 17*, (1990), pp. 211-240.

[38] Mukaidono, M., ‘Several Extensions of Truth Values in Fuzzy Logic’, *Proceedings ISMIS’99*, (June 1999), pp. 282-291.

[39] Nguyen, H. T., ‘On Fuzziness and Linguistic Probabilities’, *Journal of Mathematical Analysis and Applications, 61*, (1977), pp. 658-671.

[40] Pawlak, Z. ’Rough sets’, *International Journal Computation and Information Science, 11*, (1982), pp. 341-356.

[41] Pawlak, Z., *ROUGH SETS Theoretical Aspects of Reasoning about Data*, (Kluwer Academic Publishers, 1991).

[42] Pawlak, Z., Skowron, A., 'Rough Membership Functions', *Fuzzy Logic for the Management of Uncertainty* (L.A. Zadeh and J. Kacprzyk, Eds.), (Wiley, New York, 1994), pp. 251-271.
[43] Polkowski, L. and Skowron, A. (Eds.), *Rough Sets in Knowledge Discovery, I, II*, Physica-Verlag, Heidelberg, (1998).
[44] Richard Jeffrey, 'Probabilistic Thinking', Princeton University (1995).
[45] Shafer, G., A Mathematical Theory of Evidence, Princeton University Press, Princeton, (1976).
[46] Shenoi, S., Melton, A., 'Proximity Relations in the Fuzzy Relational Database Model', *Fuzzy Sets and Systems* 31, (1989), pp. 285-296.
[47] Slowinski, R., Vanderpooten, D., 'A Generalized Definition of Rough Approximations Based on Similarity', *IEEE Transactions on Knowledge and Data Engineering*, (2000), Vol. 12, No.2, pp. 331-336.
[48] Sugeno, M., 'Fuzzy Measures and Fuzzy Integrals-a survey', In: Gupta, Saridis and Gaines, (1977), pp. 89-102.
[49] Turksen, I.B., 'Upper and Lower Memberships in CWW', *Proceedings of FUZZ-IEEE*, (2001), pp. 778-780.
[50] Tversky, A., 'Features of Similarity', *Psychological Rev. 84(4)*, (1977), pp. 327-353.
[51] Yager, R.R., 'Ordinal Measures of Specificity', *Int. J. General Systems, Vol.17*, (1990), pp. 57-72.
[52] Yao, Y.Y., 'Two views of the theory of rough sets in finite universe', *International Journal of Approximate Reasoning 15*, (1996), pp. 291-317.
[53] Yao, Y.Y., 'Combination of rough and fuzzy sets based on α-level sets', in: *Rough Sets and Data Mining: Analysis for Imprecise Data*, (Lin, T.Y. and Cercone, N., Eds.), (Kluwer Academic Publishers, Boston, 1997), pp. 301-321.
[54] Yao, Y.Y., 'A comparative study of fuzzy sets and rough sets', *International Journal of Information Science 109*, (1998), pp. 227-242.

[55] Yao, Y.Y., Zhang, J.P., 'Interpreting Fuzzy Membership Functions in the Theory of Rough Sets', *Proceedings of RSCTC'00*, (2000), pp. 50-57.
[56] Zadeh, L.A., 'Fuzzy Sets and Systems', *International Journal of General Systems, Vol. 17,* (1990), pp. 129-138.
[57] Zadeh, L.A., 'Probability Measures of Fuzzy Events', *Journal of Mathematical Analysis and Applications, Vol. 23,* (1968), pp. 421-427.
[58] Zadeh, L.A., 'Similarity Relations and Fuzzy Orderings', *Inform. Sci. 3(2)*, (1970), pp. 177-200.
[59] Zadeh, L.A., 'Discussion: Probability Theory and Fuzzy Logic Are Complementary Rather Than Competitive', *Technometrics, Vol. 37, No.3*, (1995), pp. 271-276.
[60] Zadeh, L.A., 'Toward a Logic of Perceptions Based on Fuzzy Logic', *Discovering the World with Fuzzy Logic*, (V. Novak, I. Perfilieva, Eds.), Physica-Verlag, (2000), pp. 4-28.