Hybrid Probabilistic Models of Fuzzy and Rough Events

Rolly Intan[*,**], **Masao Mukaidono**[*] and **Hung T. Nguyen**[***]
[*]Department of Computer Science, Meiji University,
Kawasaki-shi, Kanagawa-ken, Japan
[**]Petra Christian University
Siwalankerto 121-131, Surabaya, Indonesia 60236
[***]New Mexico State University
Las Cruces, NM 88003-8001, USA
{rolly,masao}@cs.meiji.ac.jp, hunguyen@nmsu.edu

This paper discusses the relationship between probability and fuzziness based on the process of perception. As a generalization of crisp set, fuzzy set is used to model fuzzy event as proposed by Zadeh. Similarly, we may consider rough set to represent rough event in terms of probability measure. Special attention will be given to conditional probability of fuzzy event as well as conditional probability of rough event. Their several combinations of formulation and properties are examined. In the relation to evidence theory, the probability of rough event may be considered as a connecting bridge between belief-plausibility measures and the probability measures. Moreover, generalized fuzzy-rough event is introduced to generalize both fuzzy and rough events.

Keywords: Probability of Fuzzy Event, Probability of Rough Event, Evidence Theory, Generalized Fuzzy-Rough Event.

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 1977 [13], Kosko 1990 [12], Zadeh 1968 [20], 1995 [21], and so on.

In this paper, again we tried to simply understand the relationship between probability and fuzziness using the process of perception performed by human being as discussed in [5]. 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 four 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* [20]. 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 complementary tools.

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 sets represent deterministic uncertainty by a class or classes which do not possess sharply defined boundaries [19]. Randomness of fuzzy events can be quantified by probabilities, see [20]. 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 [3, 4] 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 [15], 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 [4, 11, 15, 18]. By rough set theory, we propose a rough event representing two approximate events, namely lower and upper approximate events, in the presence of probability theory providing *probability of rough event.* Therefore, a rough event might be considered as approximation of a given crisp event. Moreover, probability of a rough event gives semantic formulation of interval probability. Formulation of interval probability is useful in order to represent the worst and the best case in decision making process. In this paper, special attention will be given to conditional probability of rough events providing several combinations of formulation and properties.

In addition, a generalized fuzzy rough set as proposed in [10, 8] is an approximation of a given fuzzy set on a given fuzzy covering. Since fuzzy sets generalizes crisp sets and fuzzy covering generalizes crisp partition, the generalized fuzzy rough set is considered as the most general extension of fuzzy set and rough set as well as rough fuzzy set and fuzzy rough set as proposed in [1]. Thus, using 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, we show and discuss 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. Probability of Fuzzy Events

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 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 [20]:

Definition 1 *Let* $(U, \mathcal{F}, P)$ *be a probability space. Then, a fuzzy event* A *is a fuzzy set* A *on* U *whose membership function,* $\mu_A : U \to [0,1]$ *is* $\mathcal{F}$*-measurable. The probability of fuzzy event* A *is defined by:*
- continuous sample space:

$$P(A) = \int_U \mu_A(u)dP = \int_U \mu_A(u)p(u)du, \quad (1)$$

- discrete sample space:

$$P(A) = \sum_U \mu_A(u)p(u), \quad (2)$$

where $p(u)$ *is the probability density function of* P.

For example, given a sentence "John ate *a few* eggs for breakfast" in which we do not know exactly how many eggs John ate for breakfast. Instead, arbitrarily given probability distribution function of "John ate $u \in U$ egg(s) for breakfast" as shown in Table 1. "a few" is a fuzzy label that also means a fuzzy

Table 1: Probability Distribution of u

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

event as arbitrarily given by the following fuzzy set: $a\ few = \{1/1, 0.6/2, 0.2/3\}$, where $\mu_{afew}(2) = 0.6$. By Definition 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 [19],

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

Containment: $A \subset B \Longleftrightarrow \mu_A(u) \le \mu_B(u),\ \forall u,$

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

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

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

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

(1) $A \subset B \Longrightarrow P(A) \le P(B),$

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

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

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

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

(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 the probability of the event occurring given that another event has already occurred. Specifically,

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

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 follows: for A and B fuzzy sets on U,

$$P(A|B) = \frac{\sum_u \min\{\mu_A(u), \mu_B(u)\}p(u)}{\sum_u \mu_B(u)p(u)},\ \forall u \in U, \quad (3)$$

where $\sum_u \mu_B(u)p(u) > 0$. Some properties are: for A and B fuzzy sets on U,

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

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

$$P(A) = \sum_k P(B_k)P(A|B_k),$$

(3) Bayes Theorem:
$P(A|B) = [P(B|A) \times P(A)]/P(B).$

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 density function, $p(u) = 1/|U|$, is regarded as a constant variable. Therefore, conditional probability of fuzzy event A given B is defined simply as:

$$P(A|B) = \frac{\sum_u \min\{\mu_A(u), \mu_B(u)\}}{\sum_u \mu_B(u)}, \ \forall u \in U. \quad (4)$$

In [3, 4], we used this formula to calculate degree of similarity relationship between two fuzzy labels (sets) and called it *fuzzy conditional probability relation.*

3. Probability of Rough Events

Rough sets are another generalization of crisp sets by studying sets with imprecise boundaries. 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 [4, 11, 15, 18]. 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,

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

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

$$\begin{aligned} \overline{A} &= \{u \in U \mid [u]_R \cap A \neq \emptyset\}, \\ &= \bigcup\{[u]_R \in U/R \mid [u]_R \cap A \neq \emptyset\}, \end{aligned}$$

is the 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 approximation and generalization of a given crisp event. Probability of rough event is then defined as follows.

Definition 2 *Let $(U, \mathcal{F}, P)$ be a probability space. Then, a rough event of $A = [\underline{A}, \overline{A}] \in \mathcal{F}^2$ is a pair of lower and upper approximation of $A \subseteq U$. The probability of rough event A is defined by an interval probability $[P(\underline{A}), P(\overline{A})]$, where $P(\underline{A})$ and $P(\overline{A})$ are lower and upper probabilities, respectively.*

- lower probability:

$$\begin{aligned} P(\underline{A}) &= \sum_{\{u \in U \mid [u]_R \subseteq A\}} p(u) & (5) \\ &= \sum_{\bigcup\{[u]_R \in U/R \mid [u]_R \subseteq A\}} P([u]_R), & (6) \end{aligned}$$

- upper probability:

$$\begin{aligned} P(\overline{A}) &= \sum_{\{u \in U \mid [u]_R \cap A \neq \emptyset\}} p(u), & (7) \\ &= \sum_{\bigcup\{[u]_R \in U/R \mid [u]_R \cap A \neq \emptyset\}} P([u]_R), & (8) \end{aligned}$$

where $p(u)$ is probability density function of P.

The definition shows that the probability of a rough event gives semantic formulation of interval probability. By combining with other set-theoretic operators such as $\neg$, $\cup$ and $\cap$, we have the following results:

(P1) $P(\underline{A}) \leq P(A) \leq P(\overline{A})$,

(P2) $A \subseteq B \Leftrightarrow [P(\underline{A}) \leq P(\underline{B}), P(\overline{A} \leq P(\overline{B})]$,

(P3) $P(\underline{\neg A}) = 1 - P(\overline{A}),\ P(\overline{\neg A}) = 1 - P(\underline{A})$,

(P4) $P(\neg \underline{A}) = P(\overline{\neg A}),\ P(\neg \overline{A}) = P(\underline{\neg A})$,

(P5) $P(\underline{U}) = P(U) = P(\overline{U}) = 1$,
$P(\underline{\emptyset}) = P(\emptyset) = P(\overline{\emptyset}) = 0$,

(P6) $P(\underline{A \cap B}) = P(\underline{A} \cap \underline{B}),\ P(\overline{A \cap B}) \leq P(\overline{A} \cap \overline{B})$,

(P7) $P(\underline{A \cup B}) \geq P(\underline{A}) + P(\underline{B}) - P(\underline{A \cap B})$,

(P8) $P(\overline{A \cup B}) \leq P(\overline{A}) + P(\overline{B}) - P(\overline{A \cap B})$,

(P9) $P(A) \leq P(\underline{(\overline{A})}),\ P(A) \geq P(\overline{(\underline{A})})$,

(P10) $P(\underline{A}) = P(\underline{\underline{A}}),\ P(\overline{A}) = P(\overline{\overline{A}})$,

(P11) $P(\underline{A} \cup \underline{\neg A}) \leq 1,\ P(\overline{A} \cup \overline{\neg A}) \geq 1$,

(P12) $P(\underline{A} \cap \underline{\neg A}) = 0,\ P(\overline{A} \cap \overline{\neg A}) \geq 0$.

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

$$(1)\ P(\underline{A}|\underline{B}) = \frac{P(\underline{A} \cap \underline{B})}{P(\underline{B})}, \quad (2)\ P(\underline{A}|\overline{B}) = \frac{P(\underline{A} \cap \overline{B})}{P(\overline{B})},$$

$$(3)\ P(\overline{A}|\underline{B}) = \frac{P(\overline{A} \cap \underline{B})}{P(\underline{B})}, \quad (4)\ P(\overline{A}|\overline{B}) = \frac{P(\overline{A} \cap \overline{B})}{P(\overline{B})}.$$

Some relations are given by:

$$P(\underline{A} \cap \underline{B}) \leq P(\overline{A} \cap \underline{B}) \Rightarrow P(\underline{A}|\underline{B}) \leq P(\overline{A}|\underline{B}),$$

$$P(\underline{A} \cap \overline{B}) \leq P(\overline{A} \cap \overline{B}) \Rightarrow P(\underline{A}|\overline{B}) \leq P(\overline{A}|\overline{B}).$$

Similarly, they also satisfy some properties:

(1) Normalization:

$$(i)\ P(\underline{A}|\underline{B}) + P(\underline{\neg A}|\underline{B}) \leq 1,$$
$$(ii)\ P(\underline{A}|\overline{B}) + P(\underline{\neg A}|\overline{B}) \leq 1,$$
$$(iii)\ P(\overline{A}|\underline{B}) + P(\overline{\neg A}|\underline{B}) \geq 1,$$
$$(iv)\ P(\overline{A}|\overline{B}) + P(\overline{\neg A}|\overline{B}) \geq 1.$$

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

$$(i)\ P(\underline{A}) \geq \sum_k P(\underline{B_k})P(\underline{A}|\underline{B_k}),$$
$$(ii)\ P(\underline{A}) \leq \sum_k P(\overline{B_k})P(\underline{A}|\overline{B_k}),$$
$$(iii)\ P(\overline{A}) \geq \sum_k P(\underline{B_k})P(\overline{A}|\underline{B_k}),$$
$$(iv)\ P(\overline{A}) \leq \sum_k P(\overline{B_k})P(\overline{A}|\overline{B_k}).$$

Note: $\{B_k|k \in \mathbb{N}_n\}$ might be different from U/R.

(3) Bayes Theorem:

$$(i)\ P(\underline{A}|\underline{B}) = \frac{P(\underline{B}|\underline{A})P(\underline{A})}{P(\underline{B})},$$
$$(ii)\ P(\underline{A}|\overline{B}) = \frac{P(\overline{B}|\underline{A})P(\underline{A})}{P(\overline{B})},$$
$$(iii)\ P(\overline{A}|\underline{B}) = \frac{P(\underline{B}|\overline{A})P(\overline{A})}{P(\underline{B})},$$
$$(iv)\ P(\overline{A}|\overline{B}) = \frac{P(\overline{B}|\overline{A})P(\overline{A})}{P(\overline{B})}.$$

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

$$(1)\ P_1(A|B) = \frac{P(\underline{A \cap B})}{P(\underline{B})}, \quad (2)\ P_2(A|B) = \frac{P(\underline{A \cap B})}{P(\overline{B})},$$

$$(3)\ P_3(A|B) = \frac{P(\overline{A \cap B})}{P(\underline{B})}, \quad (4)\ P_4(A|B) = \frac{P(\overline{A \cap B})}{P(\overline{B})}.$$

Also some relations concerning the above formulations are given by:

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

They satisfy some properties of conditional probability:

(1) Normalization:

$$(i)\ P_1(A|B) + P_1(\neg A|B) \leq 1,$$
$$(ii)\ P_2(A|B) + P_2(\neg A|B) \leq 1,$$
$$(iii)\ P_3(A|B) + P_3(\neg A|B) \geq 1,$$
$$(iv)\ P_4(A|B) + P_4(\neg A|B) \geq 1.$$

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

$$(i)\ P(\underline{A}) \geq \sum_k P(\underline{B_k})P_1(A|B_k),$$
$$(ii)\ P(\underline{A}) \geq \sum_k P(\overline{B_k})P_2(A|B_k),$$
$$(iii)\ P(\overline{A}) \leq \sum_k P(\underline{B_k})P_3(A|B_k),$$
$$(iv)\ P(\overline{A}) \leq \sum_k P(\overline{B_k})P_4(A|B_k).$$

Note: $\{B_k|k \in \mathbb{N}_n\}$ might be different from U/R.

(3) Bayes Theorem:

$$(i)\ P_1(A|B) = \frac{P_1(B|A)P(\underline{A})}{P(\underline{B})},$$
$$(ii)\ P_2(A|B) = \frac{P_2(B|A)P(\underline{A})}{P(\overline{B})},$$
$$(iii)\ P_3(A|B) = \frac{P_3(B|A)P(\overline{A})}{P(\underline{B})},$$
$$(iv)\ P_4(A|B) = \frac{P_4(B|A)P(\overline{A})}{P(\overline{B})}.$$

4. Probability of Generalized Fuzzy-Rough Events

A generalized fuzzy rough set is an approximation of a given fuzzy set on a given fuzzy covering. Since fuzzy sets generalize crisp sets and covering generalizes partition, fuzzy covering is regarded as the most general 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[9].

Definition 3 *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*

$$\sum_{i=1}^{m} \mu_{C_i}(u_k) \geq 1, \quad \forall k \in \mathbb{N}_n, \tag{9}$$

$$0 < \sum_{k=1}^{n} \mu_{C_i}(u_k) < n, \quad \forall i \in \mathbb{N}_m, \tag{10}$$

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

Given a fuzzy set A on fuzzy covering as defined in Definition 3, a generalized fuzzy rough set A is defined in the following definition.

Definition 4 *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. $\underline{A}_m$, $\underline{A}_M$, $\overline{A}_m$ and $\overline{A}_M$ are defined as minimum lower, maximum lower, minimum upper and maximum upper approximate fuzzy sets of A, respectively, as follows.*

$$\mu_{\underline{A}_m}(y) = \inf_{\{i|\mu_{C_i}(y)>0\}} \min_{\{z\in U|\mu_{C_i}(z)>0\}} \{\psi(i,z)\} \tag{11}$$

$$\mu_{\underline{A}_M}(y) = \sup_{\{i|\mu_{C_i}(y)>0\}} \min_{\{z\in U|\mu_{C_i}(z)>0\}} \{\psi(i,z)\} \tag{12}$$

$$\mu_{\overline{A}_m}(y) = \inf_{\{i|\mu_{C_i}(y)>0\}} \max_{z\in U} \{\psi(i,z)\}, \tag{13}$$

$$\mu_{\overline{A}_M}(y) = \sup_{\{i|\mu_{C_i}(y)>0\}} \max_{z\in U} \{\psi(i,z)\}, \tag{14}$$

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

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

$$\underline{A}_m \subseteq \underline{A}_M \subseteq \overline{A}_M, \ \underline{A}_m \subseteq \overline{A}_m \subseteq \overline{A}_M, \ \underline{A}_M \subseteq A.$$

Iterative is applied for almost all approximate fuzzy sets except for $\underline{A}_M$ as follows.

$$\underline{A}_m \supseteq \underline{(\underline{A}_m)}_m \supseteq \cdots \supseteq \underline{A}_{m*}, \ \overline{A}_m \subseteq \overline{(\overline{A}_m)}_m \subseteq \cdots \subseteq \overline{A}_{m*},$$

$$\overline{A}_M \subseteq \overline{(\overline{A}_M)}_M \subseteq \cdots \subseteq \overline{A}_{M*},$$

where $\underline{A}_{m*}$, $\overline{A}_{m*}$ and $\overline{A}_{M*}$ are the lowest approximation of $\underline{A}_m$, the uppermost approximation of $\overline{A}_m$ and the uppermost approximation of $\overline{A}_M$, 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 5 *Let $(U, \mathcal{F}, P)$ be a probability space. Then, a generalized fuzzy-rough event of $A = [\underline{A}_m, \underline{A}_M, \overline{A}_m, \underline{A}_M] \in \mathcal{F}^4$ 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 $[P(\underline{A}_m), P(\underline{A}_M), P(\overline{A}_m), P(\underline{A}_M)]$ as follows.*

$$P(\underline{A}_m) = \sum_U \mu_{\underline{A}_m}(u)p(u), \tag{15}$$

$$P(\underline{A}_M) = \sum_U \mu_{\underline{A}_M}(u)p(u), \tag{16}$$

$$P(\overline{A}_m) = \sum_U \mu_{\overline{A}_m}(u)p(u), \tag{17}$$

$$P(\overline{A}_M) = \sum_U \mu_{\overline{A}_M}(u)p(u), \tag{18}$$

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

By combining with other set-theoretic operators such as $\neg$, $\cup$ and $\cap$, we have the following properties:

(PG1) $P(\underline{A}_m) \leq P(\underline{A}_M) \leq P(\overline{A}_M)$,
$P(\underline{A}_M) \leq P(A)$, $P(\underline{A}_m) \leq P(\overline{A}_m) \leq P(\overline{A}_M)$,

(PG2) $A \subseteq B \Leftrightarrow [P(\underline{A}_m) \leq P(\underline{B}_m), P(\underline{A}_M) \leq P(\underline{B}_M), P(\overline{A}_m) \leq P(\overline{B}_m), P(\overline{A}_M) \leq P(\overline{B}_M)]$,

(PG3) $P(\underline{U}_\lambda) \leq 1$, $P(\overline{U}_\lambda) \leq 1$, $P(\underline{\emptyset}_\lambda) = P(\overline{\emptyset}_\lambda) = 0$,

(PG4) $P(\underline{A \cap B}_\lambda) \leq P(\underline{A}_\lambda \cap \underline{B}_\lambda)$, $P(\overline{A \cap B}_\lambda) \leq P(\overline{A}_\lambda \cap \overline{B}_\lambda)$,

(PG5) $P(\underline{A \cup B}_\lambda) \geq P(\underline{A}_\lambda) + P(\underline{B}_\lambda) - P(\underline{A \cap B}_\lambda)$,

(PG6) $P(\overline{A \cup B}_\lambda) \leq P(\overline{A}_\lambda) + P(\overline{B}_\lambda) - P(\overline{A \cap B}_\lambda)$,

(PG7) $P(\underline{A}_m) \geq P(\underline{(\underline{A}_m)}_m) \geq \cdots \geq P(\underline{A}_{m*})$,
$P(\underline{A}_M) = P(\underline{(\underline{A}_M)}_M)$,

(PG8) $P(\overline{A}_\lambda) \leq P(\overline{(\overline{A}_\lambda)}) \leq \cdots \leq P(\overline{A}_{\lambda*})$,

(PG9) $P(\underline{A}_\lambda \cup \underline{\neg A}_\lambda) \leq 1$,

(PG10) $P(\underline{A}_\lambda \cap \underline{\neg A}_\lambda) \geq 0$, $P(\overline{A}_\lambda \cap \overline{\neg A}_\lambda) \geq 0$,

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

5. Belief and Plausibility Measures

Belief and plausibility measures are mutually dual functions in evidence theory originally introduced by Glenn Shafer in 1976 [16]. This work was motivated and related to lower and upper probability by Dempster in 1967 [2] in which these all types of measures are subsumed into the concept of fuzzy measure proposed by Sugeno in 1977 [17]. 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 Section 3 [5, 6]. 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 [20]. 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 [6, 7]. Belief and plausibility measures can be represented by a single function called basic probability assignment as defined by the following [11]:

Definition 6 *Let U be a sample space and $\mathcal{P}(U)$ be power set of U,*

$$m : \mathcal{P}(U) \to [0, 1] \tag{19}$$

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.

Note that:

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

$$Bel(A) = \sum_{E \subseteq A} m(E) \tag{20}$$

$$Pl(A) = \sum_{E \cap A \neq \emptyset} m(E). \tag{21}$$

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 dual functions, i.e., $Pl(A) = 1 - Bel(\neg A)$. Similarly, $Bel(A) = 1 - Pl(\neg A)$.

Since belief and plausibility measures are defined as above, 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,

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

$$Bel^{-1}(A) = \bigcup_{E \in \mathcal{P}(U), E \subseteq A} E,$$

where $Pl(A)$ and $Pl^{-1}(A)$ correspond to $P(\overline{A})$ and $\overline{A}$, respectively. Similarly, $Bel(A)$ and $Bel^{-1}(A)$ correspond to $P(\underline{A})$ and $\underline{A}$, respectively. Hence, property of $P(\overline{A}) = P(\overline{\overline{A}})$ 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, $P(A) \geq Pl(Bel^{-1}(A))$ of$P(A) \geq P((\underline{A}))$, 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*, $p : U \to [0, 1]$, 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.

In fact, belief functions, as well as possibility measures, can be rigorously formulated in terms of a powerful ingredient, *random set theory*, see [14], in which the above mentioned non-additive set-functions are related to capacity functionals of random sets, the counter-parts of distribution functions of random vectors. Specifically, the so-call "basic probability assignment m" in the Definition 6 is nothing else than the probability density function of a random set and its associate belief measure is the corresponding probability distribution of a random set.

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 ($p(u)$) (see Eq. (5) and (7)). Probability of the generalized fuzzy-rough event was proposed along with its properties.

We may then summarize their relation by the following figure.

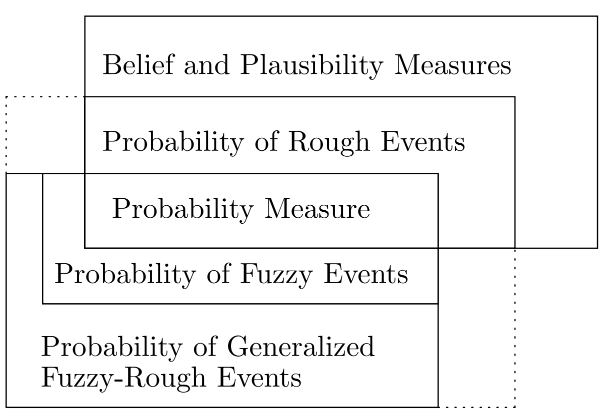

Figure 1: Generalization based on Crisp-Granularity and Membership Function

Reference:

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

2) Dempster, A.P., 'Upper and Lower Probability Induced by Multi-valued Mappings', *Annals of Mathematical Statistics, 38,* (1967), pp.325-339.

3) Intan, R., Mukaidono, M., 'Conditional Probability Relations in Fuzzy Relational Database ', *Proceedings of RSCTC'00, LNAI 2005, Springer & Verlag*, (2000), pp.251-260.

4) Intan, R., Mukaidono, M.,Yao, Y.Y., 'Generalization of Rough Sets with α-coverings of the Universe Induced by Conditional Probability Relations', *Proceedings of International Workshop on Rough Sets and Granular Computing, LNAI 2253, Springer & Verlag*, (2001), pp.311-315.

5) Intan, R., Mukaidono, M., 'A Proposal of Probability of Rough Event based on Probability of Fuzzy Event', *Proceeding of The 3nd International Conference on Rough Sets and Current Trend In Computing 2002, LNAI 2475, Springer & Verlag*, (2002), pp. 357-364.

6) Intan, R., Mukaidono, M., 'Probability of Fuzzy Event to Probability of Rough Event', *Proceeding of Fuzzy Sets and Knowledge Discovery 2002*, (2002), pp. 549-553.

7) Intan, R., Mukaidono, M., 'From Evidence Theory to Probability of Generalized Fuzzy-Rough Events', *Proceeding of Fuzzy Information Processing 2003*, (2003), pp. 883-888.

8) Intan, R., Mukaidono, M., 'Generalization of Fuzzy Rough Sets By Fuzzy Covering Based On Weak Fuzzy Similarity Relation', *Proceeding of Fuzzy Sets and Knowledge Discovery 2002*, (2002), pp. 344-348.

9) Intan, R., Mukaidono, M., 'Degree of Similarity in Fuzzy Partition', *Proceedings of AFSS'02, LNAI 2275, Springer & Verlag*, (2002), pp. 20-26.

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

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

12) Kosko, B., 'Fuzziness VS. Probability', *International Journal of General Systems, Vol 17*, pp. 211-240.

13) Nguyen, H. T., 'On Fuzziness and Linguistic Probabilities', *Journal of Mathematical Analysis and Applications, 61*, pp. 658-671.

14) Nguyen, H.T., Bouchon-Meunier, B., ‘Random Sets and Large Deviations Principle as a Foundation for Possibility Measures’, (2003), to appear in *Journal of Soft Computing.*

15) Pawlak, Z., ‘Rough sets’, *International Journal Computation & Information Science*, **11**, (1982), pp. 341-356.

16) Shafer, G., *A Mathematical Theory of Evidence, Princeton University Press, Princeton,* (1976).

17) Sugeno, M., ‘Fuzzy Measures and Fuzzy Integrals-a survey’, *In: Gupta, Saridis, and Gaines,* (1977), pp. 89-102.

18) Yao, Y.Y., ‘Two views of the theory of rough sets in finite universe’, *International Journal of Approximate Reasoning 15*, (1996), pp.291-317.

19) Zadeh, L.A., ‘Fuzzy Sets and Systems’, *International Journal of General Systems, Vol 17*, (1990), pp. 129-138.

20) Zadeh, L.A., ‘Probability Measures of Fuzzy Events’, *Journal of Mathematical Analysis and Applications, Vol 23*, (1968), pp. 421-427.

21) Zadeh, L.A., ‘Discussion: Probability Theory and Fuzzy Logic Are Complementary Rather Than Competitive’, *Technometrics, Vol 37, No.3*, (1995), pp. 271-276.