On the Relation of Probability, Fuzziness, Rough and Evidence Theory

Rolly Intan

Petra Christian University
Department of Informatics Engineering
Surabaya, Indonesia
rintan@petra.ac.id

Abstract. Since the appearance of the first paper on fuzzy sets proposed by Zadeh in 1965, the relationship between probability and fuzziness in the representation of uncertainty has been discussed among many people. The question is whether probability theory itself is sufficient to deal with uncertainty. In this paper the relationship between probability and fuzziness is analyzed by the process of perception to simply understand the relationship between them. It is clear that probability and fuzziness work in different areas of uncertainty. Here, fuzzy event in the presence of probability theory provides *probability of fuzzy event* in which fuzzy event could be regarded as a generalization of crisp event. Moreover, in rough set theory, a rough event is proposed representing two approximate events, namely lower approximate event and upper approximate event. Similarly, in the presence of probability theory, rough event can be extended to be *probability of rough event*. Finally, the paper shows and discusses relation among lower-upper approximate probability (probability of rough events), belief-plausibility measures (evidence theory), classical probability measures, probability of generalized fuzzy-rough events and probability of fuzzy events.

Keywords: Probability, Rough Sets, Fuzzy Sets, Evidence Theory.

1 Introduction

Since the appearance of the first paper on fuzzy sets proposed by Zadeh in 1965, the relationship between probability and fuzziness in the representation of uncertainty has been discussed among many people. The main problem is whether probability theory itself is sufficient to deal with uncertainty. This issue has been widely discussed in many papers and written by Nguyen [15], Kosko [14] Zadeh [20], [21] and so on.

In this work, again just try to understand the relationship between probability and fuzziness using the process of perception by humans. In the process of perception, the subject (human, computer, robot, etc.) tries to recognize and describe a given object (anything, like human, plant, animal, event, condition, etc.).

To conduct a successful process of perception, subject requires adequate knowledge. On the other hand, object requires a clear definition. However, human (as subject) do not know what happens in the future and has also limited knowledge. In other words, humans

R. Intan et al. (Eds.): ICSIIT 2015, CCIS 516, pp. 3–16, 2015.
DOI: 10.1007/978-3-662-46742-8_1

are not omniscient. In this case, the subject is in a non-deterministic situation in performing a perception. On the other hand, most objects (shape, feeling, emotion, etc.) cannot generally be clearly defined. Therefore, the perception process is in uncertain situation.

To summarize the relationship between subject and object in the process of perception, there are four possible situations as follows [10].

- If the subject has sufficient knowledge and the object has a clear definition, it becomes a certainty.
- If the subject has sufficient knowledge and object has unclear definition, it comes to the situation of fuzziness. In general, fuzziness, also called deterministic uncertainty, may occur in the situation when one is able to subjectively determine or describe a given object, but somehow the object does not have a specific or clear definition. For example, a man describes a woman as a beautiful woman. Obviously definition of a beautiful woman is unclear, uncertain and subjective. The man, however is convinced of what he describes someone as a pretty woman.
- If the subject is not in having sufficient knowledge and object has a clear definition, it comes to the situation of randomness. Randomness is usually called non-deterministic uncertainty because subject cannot determine or describe a given object clearly although the object has a clear definition. Here, probability theory was developed for dealing with the random experiment. For example, in throwing a dice, even though there are six possible defined result of outcomes, one cannot ensure outcome of dice. Another example, in solving multiple choice problem, because of his limited knowledge, a student may not be assured to choose an answer out of 4 possible answers.
- If the subject is in insufficient knowledge and object definition is unclear, it comes to be a probability of fuzzy event [20]. In this situation, both the probability and fuzziness are combined. For example, how to predict an ill-defined event: "Tomorrow will be a warm day." Speaking of morning is talk about a future in which the subject cannot determine what happens in the future. The situation must be addressed by probability. However, "hot" is a subjectively defined event (called fuzzy event). Therefore, the event is regarded to be a probability of fuzzy event.

From these four situations obviously seen that the probability and the fuzziness in work in different areas of uncertainty. Probability theory itself is not sufficient to deal with ill-defined event or fuzzy event. Instead, probability and Fuzziness should be considered as a complementary tool.

In probability, set theory is used to provide a language for modeling and describing random experiments. In (classical) theory of sets, subsets of the sample space of an experiment are referred to as crisp events. Fuzzy set theory proposed by Zadeh in 1965, is regarded as a generalization of (classical) set theory in which fuzzy sets is to represent deterministic uncertainty by a class or classes that do not have sharp defined boundary [21].

In fuzzy set theory, an ill-defined event, called fuzzy event can be described in the presence of probability theory called probability of fuzzy event [20] in which fuzzy event could be regarded as a generalization of crisp event. Conditional probability as an important property in the theory of probability usually used in generating inference rule can be extended to the conditional probability of fuzzy event. In the situation of uniform distribution probability, conditional probability of fuzzy event can be simplified to be what we call fuzzy conditional probability relations as proposed in [3] and [4] to calculate the similarity of two fuzzy labels (sets).

Similarly, rough sets theory generalizes the classical set theory by studying sets with imprecise boundaries. A rough set [16], which is characterized by a pair of approximations, can be seen as an approximate representation of a given classical set in terms of two subsets derived from a partition in the universe as proposed in [12], [13], [16] and [19]. By the theory of rough sets, rough event which is consists of lower approximate event and upper approximate event, in the presence of probability theory provides probability of rough event. Therefore, rough event could be also considered as an approximation of a given crisp event. Moreover, the probability of rough event presents a semantic formulation to what called the interval probability. Formulation of the interval probability is useful to represent the worst and the best probabilities of an event for supporting decision making process. In this paper, special attention is focused to discuss conditional probability of rough event and proved that it satisfied some properties.

In addition, a generalized fuzzy rough set as proposed in [5] and [7] is considered as an approximation of a given fuzzy set on a given fuzzy covering. Since fuzzy set and fuzzy covering generalize crisp set and crisp partition respectively, the generalized fuzzy rough set is regarded as a generalization of rough fuzzy sets and rough fuzzy sets as proposed by Dubois and Prade in [2]. Therefore, a generalized fuzzy rough event represented by generalized fuzzy rough set, in the presence of probability, provides probability of generalized fuzzy rough event. The generalized fuzzy-rough event is represented in four approximates, namely lower maximum of fuzzy event, lower minimum of fuzzy event, upper maximum of fuzzy event and upper minimum of fuzzy event.

Finally, we show and discuss relation among lower-upper approximate probability (probability of rough events), belief-plausibility measures (evidence theory), classical probability measures, probability of generalized fuzzy-rough events and probability of fuzzy events.

2 Probability of Fuzzy Event

Probability theory is based on the paradigm of a random experiment in which outcome of the experiment cannot be predicted with certainty, before running the experiment. In other words, as discussed in the previous section, in the situation of probability, subject does not have sufficient knowledge to determine outcome of the 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 the universal set. In (classical) theory of sets, subsets of the sample space of an experiment are used to represent crisp events.

To represent an ill-defined event, crisp event should be generalized to the fuzzy event in which the fuzzy sets used to represent fuzzy event. Formally, probability of fuzzy event is showed in the following definition [20].

Definition 1. Let (U,F,P) be regarded as a probability space in which U is the sample space, F presents sigma algebra of events and P is a probability of an event over U. Then, a fuzzy event $A \in F$ is represented by a fuzzy set A on U whose membership function given by $\mu_A : U \to [0, 1]$. The probability of fuzzy event A is defined by the following equations:

– continuous sample space:

$$P(A) = \int_U \mu_A(u)dP = \int_U \mu_A(u) \cdot p(u)du \tag{1}$$

– discrete sample space:

$$P(A) = \sum_U \mu_A(u) \cdot p(u) \tag{2}$$

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

For example, arbitrarily given a sentence "John ate *a few* eggs for breakfast". Here, we have insufficient knowledge to know exactly how many eggs John ate for breakfast. Instead, probability distribution function of "John ate $u \in U$ egg(s) for breakfast" is arbitrarily shown by Table 1.

Table 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 meaningful fuzzy label, "*a few*" represented by a fuzzy event is arbitrarily given by a fuzzy set, $\mu_{afew} = \{1/1, 0.6/2, 0.2/3\}$, where $\mu_{afew}(2) = 0.6$. Probability of "John ate *a few* eggs for breakfast", denoted by P (*a few*), is calculated as follow.

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

Some basic concepts and operations relating to the fuzzy sets are given by the following operations. Given A and B are two fuzzy sets on U [21],

- Union: $\mu_{A\cup B}(u) = \max[\mu_A(u), \mu_B(u)]$,
- Complement: $B = \neg A \Leftrightarrow \mu_B(u) = 1 - \mu_A(u), \forall u$,
- Intersection: $\mu_{A\cap B}(u) = \min[\mu_A(u), \mu_B(u)]$,
- Sum: $\mu_{A\oplus B}(u) = \mu_A(u) + \mu_B(u) - \mu_A(u) \cdot \mu_B(u)$.
- Equality: $A = B \Leftrightarrow \mu_A(u) = \mu_B(u), \forall u$,
- Containment: $A \subset B \Leftrightarrow \mu_A(u) \le \mu_B(u), \forall u$,
- Product: $\mu_{AB}(u) = \mu_A(u) \cdot \mu_B(u)$,

It can be verified obviously that the probability of fuzzy event satisfies certain properties. Given A and B are two fuzzy sets on U,

1. $P(A \oplus B) = P(A) + P(B) - P(A \cdot B)$,
2. $A \subset B \Rightarrow P(A) \le P(B)$,
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,
4. $P(A \cap \neg A) \ge 0$.
5. $P(A \cup \neg A) \le 1$,

1, 2 and 3 prove that probability of fuzzy event satisfies additivity axiom of sum, monotonicity and additivity axiom of union, respectively. However, it does not satisfy law of non-contradiction and law of excluded middle as clearly seen in (4) and (5).

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

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

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

In discrete sample space, the conditional probability of fuzzy event could be defined as follows. Given A and B are two fuzzy sets on U,

$$P(A|B) = \frac{\sum_U \min\left[\mu_A(u), \mu_B(u)\right] \cdot p(u)}{\sum_U \mu_B(u) \cdot p(u)}, \forall u \in U, \tag{3}$$

where $\sum_U \mu_B(u) \cdot p(u) > 0$. It can be proved that several properties are satisfied in the conditional probability of fuzzy event. Given A and B be two fuzzy sets on U,

1. Normalization: $\mathrm{P}(A|B) + \mathrm{P}(\neg A|B) \geq 1$,
2. Total Probability: If $\{B_k \mid k \in \mathrm{N}_n\}$ are pairwise disjoint, crisp and exhaustive events, i.e., $\mathrm{P}(B_i \cap B_j) = 0$ for $i \neq j$ and $\cup B_k = U$, then:

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

3. Bayes Theorem:

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

Furthermore, the relationship between A and B in conditional probability of fuzzy event can be represented in three conditions as follows.

- Negative correlation:

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

- Positive correlation:

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

- Independent correlation:

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

In the situation of uniform distribution, the probability distribution function $p(u) = 1 / |U|$, is considered as a constant variable. Therefore, the conditional probability of A given B is more simply defined by eliminating $p(u)$ as given by:

$$P(A|B) = \frac{\sum_U \min\left[\mu_A(u), \mu_B(u)\right]}{\sum_U \mu_B(u)}, \forall u \in U, \tag{4}$$

In [3] and [4], the Equation (4) called *fuzzy conditional probability relation* is used to calculate degree of similarity relationship between two fuzzy labels (sets).

3 Probability of Rough Event

Rough set is considered as a generalization of crisp set by studying sets with imprecise boundaries. A rough set, characterized by a pair approximations, called lower approximation and upper approximation, can be seen as an approximate representation of a given crisp set in terms of two subsets derived from a partition in the universe as explained in [12], [13], [16] and [19].

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.

Formally, rough sets may be defined precisely as follows. Let *U* denotes a non-empty and finite universe. R is an equivalence relation on U. The equivalence relation *R* induces a partition of the universe. The partition is also known as the quotient set, and it is denoted by *U/R*. Given $[u]_R$ is the equivalence class in *U/R* that consists of u ∈ U. A rough set of a subset $A \subseteq U$ is represented by a pair of lower and upper approximations as given by the following equations. The lower approximation,

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

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

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

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

Definition 2. Let (*U,F,P*) be regarded as a probability space in which *U* is the sample space. *F* represents 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 given by a pair of approximations, called lower approximation 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 probability and upper probability, respectively.

– Lower probability:

$$P(Lo(A)) = \sum_{\{u \in U | [u]_R \subseteq A\}} p(u) = \sum_{\cup\{[u]_R \in U/R | [u]_R \subseteq A\}} P([u]_R), \tag{5}$$

– Upper probability:

$$P(Up(A)) = \sum_{\{u \in U | [u]_R \cap A \neq \emptyset\}} p(u) = \sum_{\cup\{[u]_R \in U/R | [u]_R \cap A \neq \emptyset\}} P([u]_R), \tag{6}$$

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

Lower and upper probabilities in Definition 2 can be regarded as an interval probability. By combining with other set-theoretic operators such as ¬, ∪ and ∩, we have the following results of properties:

1. $P(Lo(A)) \leq P(A) \leq P(Up(A))$,
2. $A \subseteq B \Leftrightarrow [P(Lo(A)) \leq P(Lo(B)), P(Up(A) \leq P(Up(B))]$,
3. $\mathrm{P}(Lo(\neg A)) = 1 - \mathrm{P}(Lo(A))$, $\mathrm{P}(Up(\neg A)) = 1 - \mathrm{P}(Up(A))$,
4. $\mathrm{P}(\neg Lo(A)) = \mathrm{P}(Up(\neg A))$, $\mathrm{P}(\neg Up(A)) = \mathrm{P}(Lo(\neg A))$,
5. $\mathrm{P}(Lo(U)) = \mathrm{P}(U) = \mathrm{P}(Up(U)) = 1$, $\mathrm{P}(Lo(\emptyset)) = \mathrm{P}(\emptyset) = \mathrm{P}(Up(\emptyset)) = 0$,
6. $\mathrm{P}(Lo(A \cap B)) = \mathrm{P}(Lo(A) \cap Lo(B))$, $\mathrm{P}(Up(A \cap B)) \leq \mathrm{P}(Up(A) \cap Up(B))$,
7. $\mathrm{P}(Lo(A \cup B)) \geq \mathrm{P}(Lo(A)) + \mathrm{P}(Lo(B)) - \mathrm{P}(Lo(A \cap B))$,
8. $\mathrm{P}(Up(A \cup B)) \leq \mathrm{P}(Up(A)) + \mathrm{P}(Up(B)) - \mathrm{P}(Up(A \cap B))$,
9. $\mathrm{P}(\mathrm{A}) \leq \mathrm{P}(Lo(Up(A)))$, $\mathrm{P}(\mathrm{A}) \geq \mathrm{P}(Up(Lo(A)))$,
10. $\mathrm{P}(Lo(A)) = \mathrm{P}(Lo(Lo(A)))$, $\mathrm{P}(Up(A)) = \mathrm{P}(Up(Up(A)))$,
11. $\mathrm{P}(Lo(A) \cup Lo(\neg A)) \leq 1$, $\mathrm{P}(Up(A) \cup Up(\neg A)) \geq 1$,
12. $\mathrm{P}(Lo(A) \cap Lo(\neg A)) = 0$, $\mathrm{P}(Up(A) \cap Up(\neg A)) \geq 0$.

Conditional probability of rough event could be considered in the following four combinations of formulations: Given A, B ⊆ U, conditional probability of A given B is given by

$$P(Lo(A)|Lo(B)) = \frac{P(Lo(A) \cap Lo(B))}{P(Lo(B))}, \tag{7}$$

$$P(Lo(A)|Up(B)) = \frac{P(Lo(A) \cap Up(B))}{P(Up(B))}, \tag{8}$$

$$P(Up(A)|Lo(B)) = \frac{P(Up(A) \cap Lo(B))}{P(Lo(B))}, \tag{9}$$

$$P(Up(A)|Up(B)) = \frac{P\big(Up(A) \cap Up(B)\big)}{P(Up(B))}. \tag{10}$$

It can be proved that the equations are also satisfied some relations as given by:

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

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

Similarly, it can also be verified that they satisfy some properties:

1. Normalization:

   - $\mathrm{P}(\mathrm{Lo(A)}|\mathrm{Lo(B)}) + \mathrm{P}(\mathrm{Lo}(\neg \mathrm{A})|\mathrm{Lo(B)}) \leq 1$,
   - $P(Lo(A)|Up(B)) + P(Lo(\neg A)|Up(B)) \leq 1$,
   - $P(Up(A)|Lo(B)) + P(Up(\neg A)|Lo(B)) \geq 1$,
   - $P(Up(A)|Up(B)) + P(Up(\neg A)|Up(B)) \geq 1$.

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

- $P(Lo(A)) \geq \sum_k P(Lo(B_k)) \cdot P(Lo(A)|Lo(B_k))$,
- $P(Lo(A)) \leq \sum_k P(Up(B_k)) \cdot P(Lo(A)|Up(B_k))$,
- $P(Up(A)) \geq \sum_k P(Lo(B_k)) \cdot P(Up(A)|Lo(B_k))$,
- $P(Up(A)) \leq \sum_k P(Up(B_k)) \cdot P(Up(A)|Up(B_k))$.

3. Bayes Theorem:

$$P(Lo(A)|Lo(B)) = \frac{P(Lo(B)|Lo(B)) \cdot P(Lo(A))}{P(Lo(B))},$$

$$P(Lo(A)|Up(B)) = \frac{P(Up(B)|Lo(A)) \cdot P(Lo(A))}{P(Up(B))},$$

$$P(Up(A)|Lo(B)) = \frac{P(Lo(B)|Up(A)) \cdot P(Up(A))}{P(Lo(B))},$$

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

We may also consider other definitions of conditional probability of rough event as given by the following equations. Given $A, B \subseteq U$, conditional probability of A given B can be also defined by

$$P_1(A|B) = \frac{P(Lo(A \cap B))}{P(Lo(B))}, \tag{11}$$

$$P_2(A|B) = \frac{P(Lo(A \cap B))}{P(Up(B))}, \tag{12}$$

$$P_3(A|B) = \frac{P(Up(A \cap B))}{P(Lo(B))}, \tag{13}$$

$$P_4(A|B) = \frac{P(Up(A \cap B))}{P(Up(B))}. \tag{14}$$

Also, it can be proved that the above equations satisfy some relations as follows.

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

Moreover, they also satisfy some properties of conditional probability:

1. Normalization:

- $P_1(A|B) + P_1(\neg A|B) \leq 1$,
- $P_2(A|B) + P_2(\neg A|B) \leq 1$,
- $P_3(A|B) + P_3(\neg A|B) \geq 1$,
- $P_4(A|B) + P_4(\neg A|B) \geq 1$.

2. Total Probability If $\{B_k \mid k \in \mathbb{N}_n\}$ are pairwise disjoint, crisp and exhaustive events,

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

– $P(Lo(A)) \geq \sum_k P\big(Lo(B_k)\big) \cdot P_1(A|B_k)$,
– $P(Lo(A)) \geq \sum_k P\big(Lo(B_k)\big) \cdot P_2(A|B_k)$,
– $P(Lo(A)) \leq \sum_k P\big(Lo(B_k)\big) \cdot P_3(A|B_k)$,
– $P\big(Lo(A)\big) \leq \sum_k P\big(Lo(B_k)\big) \cdot P_4(A|B_k)$.

3. Bayes Theorem:

$$P_1(A|B) = \frac{P_1(B|A) \cdot P(Lo(A))}{P(Lo(B))},$$

$$P_2(A|B) = \frac{P_2(B|A) \cdot P(Lo(A))}{P(Up(B))},$$

$$P_3(A|B) = \frac{P_3(B|A) \cdot P(Up(A))}{P(Lo(B))},$$

$$P_4(A|B) = \frac{P_4(B|A) \cdot P(Up(A))}{P(Up(B))}.$$

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 [6].

A generalized fuzzy rough set is an approximation of a given fuzzy set in a given fuzzy covering. Since fuzzy set and covering generalized crisp set and partition, respectively, fuzzy covering is considered the most generalized approximation space. Fuzzy covering can be considered as a case of fuzzy granularity. In this case, similarity classes as used in constructing the covering are considered as fuzzy sets. The following definition shows an alternative definition in constructing a fuzzy covering [6].

Definition 3. Let $U = \{u_1, ..., u_n\}$ be regarded as a universe. A fuzzy covering of U is defined by a family of fuzzy subsets or fuzzy classes of C, denoted by $C = \{C_1, C_2, ..., C_m\}$, and it satisfies

$$\sum_{i=1}^{m} \mu_{C_i}(u_k) \geq 1, \forall k \in N_n \quad (15)$$

$$0 < \sum_{k=i}^{n} \mu_{C_i}(u_k) < n, \forall i \in N_m \quad (16)$$

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

Let A be a fuzzy set on fuzzy covering as defined in Definition 3. A generalized fuzzy rough set A is then defined by the following definition.

Definition 4. Let U be regarded as a non-empty universe. $C = \{C_1, C_2, ..., C_m\}$ is a fuzzy covering on U. Given A be a fuzzy set on U, $Lo(A)_M$, $Lo(A)_m$, $Up(A)_M$ and $Up(A)_m$ are defined as minimum lower approximate, maximum lower approximate, minimum upper approximate and maximum upper approximate fuzzy set of A, respectively, as follows.

$$\mu_{Lo(A)_m}(y) = \inf_{\{i|\mu_{C_i}(y)>0\}} \inf_{\{z\in U|\mu_{C_i}(z)>0\}}\{\Psi(i,z)\}, \tag{17}$$

$$\mu_{Lo(A)_M}(y) = \sup_{\{i|\mu_{C_i}(y)>0\}} \inf_{\{z\in U|\mu_{C_i}(z)>0\}}\{\Psi(i,z)\}, \tag{18}$$

$$\mu_{Up(A)_m}(y) = \inf_{\{i|\mu_{C_i}(y)>0\}} \sup_{\{z\in U\}}\{\Psi(i,z)\}, \tag{19}$$

$$\mu_{Up(A)_M}(y) = \sup_{\{i|\mu_{C_i}(y)>0\}} \sup_{\{z\in U\}}\{\Psi(i,z)\}, \tag{20}$$

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. It can be proved that relationship among these approximations satisfy a partial order as follows.

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

The property of iterative is also applied for almost all approximate fuzzy sets except for $Lo(A)_M$ as shown in the following relations.

- $Lo(A)_{m*} \subseteq \cdots \subseteq Lo(Lo(A)_m)_m \subseteq Lo(A)_m,$
- $Up(A)_m \subseteq Up(Up(A)_m)_m \subseteq \cdots \subseteq Up(A)_{m*},$
- $Up(A)_M \subseteq Up(Up(A)_M)_M \subseteq \cdots \subseteq Up(A)_{M*},$

where $Lo(A)_{m*}$, $Up(A)_{m*}$ and $Up(A)_{M*}$ are regarded as the lowest approximation of $Lo(A)_m$, the uppermost approximation of $Up(A)_m$ and the uppermost approximation of $Up(A)_M$, respectively. Thus, in our concept of the generalized fuzzy rough set, a given fuzzy event can be represented into four fuzzy events, called *generalized fuzzy-rough event*. In the relation to probability theory, probability of generalized fuzzy-rough event is then defined in Definition 5 as follows.

Definition 5. Let (*U, F, P*) be regarded as a probability space in which *U* is the sample space, *F* is defined as 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 considered as fuzzy approximate events of *A*, where *A* is a given fuzzy set on *U* . The probability of generalized fuzzy-rough event *A* is then defined by a quadruplet $[P(Lo(A)_m), P(Lo(A)_M), P(Up(A)_m), P(Up(A)_M)]$ as given by the following equations.

$$P(Lo(A)_m) = \sum_U \mu_{Lo(A)_m}(u) \cdot p(u), \tag{21}$$

$$P(Lo(A)_M) = \sum_U \mu_{Lo(A)_M}(u) \cdot p(u), \tag{22}$$

$$P(Up(A)_m) = \sum_U \mu_{Up(A)_m}(u) \cdot p(u), \tag{23}$$

$$P(Up(A)_M) = \sum_U \mu_{Up(A)_M}(u) \cdot p(u), \tag{24}$$

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

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

- $P(Lo(A)_m) \leq P(Lo(A)_M) \leq P(Up(A)_M)$,
- $P(Lo(A)_M) \leq P(A)$,
- $P(Lo(A)_m) \leq P(Up(A)_m) \leq P(Up(A)_M)$,
- $A \subseteq B \Rightarrow [P(Lo(A)_m) \leq P(Lo(B)_m)$, $P(Lo(A)_M) \leq P(Lo(B)_M)$, $P(Up(A)_m) \leq P(Up(B)_m)$, $P(Up(A)_M) \leq P(Up(B)_M)]$,
- $P(Lo(U)_\lambda) \leq 1$, $P(Up(U)_\lambda) \leq 1$,
- $P(Lo(\emptyset)_\lambda) = P(Up(\emptyset)_\lambda) = 0$,
- $P(Lo(A \cap B)_\lambda) \leq P(Lo(A)_\lambda \cap Lo(B)_\lambda)$,
- $P(Up(A \cap B)_\lambda) \leq P(Up(A)_\lambda \cap Up(B)_\lambda)$,
- $P(Lo(A \cup B)_\lambda) \geq P(Lo(A)_\lambda) + P(Lo(B)_\lambda)) - P(Lo(A \cap B)_\lambda)$,
- $P(Up(A \cup B)_\lambda) \leq P(Up(A)_\lambda) + P(Up(B)_\lambda)) - P(Up(A \cap B)_\lambda)$,
- $P(Lo(A)_{m^*}) \leq \cdots \leq P(Lo(Lo(A)_m)_m) \leq P(Lo(A)_m)$,
- $P(Lo(A)_M) = P(Lo(Lo(A)_M)_M)$,
- $P(Up(A)_m) \leq P(Up(Up(A)_m)_m) \leq \cdots \leq P(Up(A)_{m^*})$,
- $P(Up(A)_M) \leq P(Up(Up(A)_M)_M) \leq \cdots \leq P(Up(A)_{M^*})$,
- $P(Lo(A)_\lambda \cup Lo(\neg A)_\lambda) \leq 1$,
- $P(Lo(A)_\lambda \cap Lo(\neg A)_\lambda) \geq 0$,
- $P(Up(A)_\lambda \cap Up(\neg A)_\lambda) \geq 0$,

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

5 Belief and Plausibility Measures

In evidence theory, belief and plausibility measures originally introduced by Glenn Shafer in 1976 [17] are mutually dual functions in evidence theory. This concept was strongly motivated and related to lower probability and upper probability proposed by Dempster in 1967 [1] in which all measures are subsumed in the concept of fuzzy measure proposed by Sugeno in 1977 [18]. Belief and plausibility Measures can be represented by a single function, called basic probability assignment, providing evidence grades for specific subsets of the universal set. In a special case when the subsets of the universal set are disjoint and each subset represents elementary set of indiscernible space, we can consider belief and plausibility measures as lower approximate probability and upper approximate probability in terms of probability of rough events as proposed in [8] and [9]. Here, lower approximate probability and upper approximate probability are considered as a special case of belief and plausibility measures, respectively in which probability of elementary set is considered as a special case of basic probability assignment. In other words, belief and plausibility measures are based on crisp-granularity in terms of a covering. On the other hand, lower approximate probability and upper approximate probability are defined on crisp-granularity in terms of disjoint partition. Here, 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. Therefore, it can be verified that lower approximate probability and upper approximate probability of a given rough event will be clearly reduced into a single value of probability. Lower approximate probability and upper approximate probability as well as belief and plausibility measures are also regarded as generalization of probability measures in the presence of crisp granularity of sample space. We may consider another generalization when the membership degree of every element of sample space in representing an event is regarded from 0 to 1. The generalization provides a concept called probability measures of fuzzy events as proposed by Zadeh in 1968 [20]. Moreover, it may then propose a more generalized probability measures by given a fuzzy event in the presence of fuzzy-granularity of sample space. The generalization called probability measures of generalized fuzzy-rough events as already discussed in previous section [5] and [7]. Belief and plausibility measures can be also represented by a single function called basic probability assignment as defined by the following [11]:

Definition 6. Given U be regarded as a universal sample space and $\mathcal{P}(U)$ be a power set of U,

$$m : \mathcal{P}(\underline{U}) \rightarrow [0, 1] \tag{25}$$

such that $\sum_{E\in\mathcal{P}(U)} m(E) = 1$ and $m(\emptyset) = 0$, 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 necessary that $m(U) = 1$.
2. It is not necessary 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 $m(E) > 0$. It is possible that focal elements may take overlap one to each other. Belief and Plausibility measures are then given by Equation (26) and (27), respectively. For $A \in \mathcal{P}(U)$,

$$Bel(A) = \sum_{E\subseteq A} m(E), \tag{26}$$

$$Pl(A) = \sum_{E\cap A\neq\emptyset} m(E). \tag{27}$$

It can be also verified that for all $A \in \mathcal{P}(U)$, $Bel(A) \leq Pl(A)$. As mention before, it can be proved that belief and plausibility measures are mutually dual functions as shown in the following equations.

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

Similarly,

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

Belief and plausibility measures are defined on a covering. Therefore, some properties are not satisfied, especially iterative properties of lower approximate probability and upper approximate probability such as given by $P(A) \leq P(Lo(Up(A)))$, $P(A) \geq P(Up(Lo(A)))$ and $P(Lo(A)) = P(Lo(Lo(A)))$, $P(Up(A)) = P(Up(Up(A)))$ as also mentioned in Section 3. 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 $P(Up(A)) = P(Up(Up(A)))$ can be represented as $Pl(A) = Pl(Pl^{-1}(A))$ by using expression of plausibility measures. It can be easily verified that the property is not satisfied instead $Pl(A) \leq Pl(P\Gamma^{-1}(A))$. Also, $P(A) \geq Pl(Bel\Gamma^{-1}(A))$ in the relation to $P(A) \geq P(Up(Lo(A)))$, 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 approximate probability and upper approximate probability fuse into a single value of probability in which probability satisfies additivity axiom as an intersection area between supperadditive property of lower approximate probability and subadditive property of upper approximate probability as respectively given by the equations, $\mathrm{P}(Lo(A \cup B)) \geq \mathrm{P}(Lo(A)) + \mathrm{P}(Lo(B)) - \mathrm{P}(Lo(A \cap B))$ and $\mathrm{P}(Up(A \cup B)) \leq \mathrm{P}(Up(A)) + \mathrm{P}(Up(B)) - \mathrm{P}(Up(A \cap B))$ as also already mentioned in Section 3.

6 Conclusion

The relationship between probability and fuzziness was discussed clearly based on the process of perception. Probability and fuzziness work in different areas of uncertainty; therefore, probability theory itself is not sufficient to deal with the uncertainty in real world application. Instead, fuzziness and probability must be regarded as a complementary concepts to represent various type of uncertainty. For instance, relation between fuzziness and probability may provide a concept, called probability of fuzzy event in which fuzzy event was represented by a given fuzzy set. Here, fuzzy event and fuzzy set are considered as a generalization of crisp event and crisp set, respectively. Similarly, rough set, as another generalization of crisp set, is used to represent rough event. In the presence of probability theory, probability of rough event was also proposed as another generalization of probability measure. Conditional probability of fuzzy event and conditional probability of rough event were examined together with their some properties. A more generalized fuzzy rough set is then proposed as an approximation of a given fuzzy set in a given fuzzy covering.

Therefore, using the concept of generalized fuzzy rough set, a generalized fuzzy-rough event was proposed as the most generalization of fuzzy event as well as rough event. Probability of the generalized fuzzy-rough event was introduced along with its properties. Figure 1 shows summary of relationship among all the concepts.

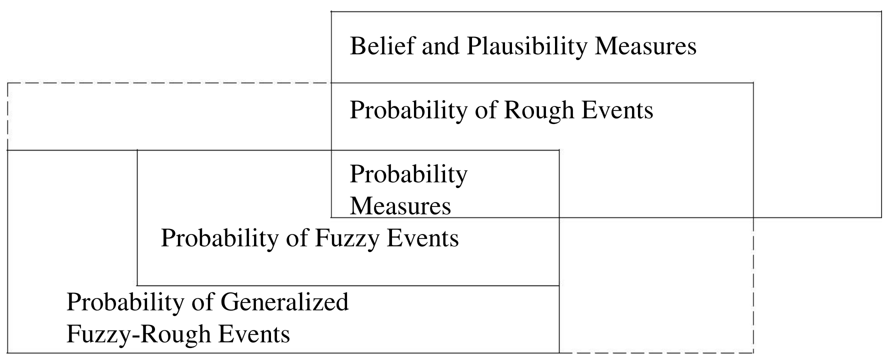

Fig. 1. Generalization based on Crisp-Granularity and Membership Function

References

1. Dempster, A.P.: Upper and Lower Probability Induced by Multivalued Mappings. Annals of Mathematical Statistics 38, 325–339 (1967)
2. Dubois, D., Prade, H.: Rough Fuzzy Sets and Fuzzy Rough Set. International Journal of General Systems 17, 191–209 (1990)
3. Intan, R., Mukaidono, M.: Application of Conditional Probability in Constructing Fuzzy Functional Dependency (FFD). In: Proceedings of AFSS 2000, pp. 271–276 (2000a)
4. Intan, R., Mukaidono, M.: Fuzzy Functional Dependency and Its Application to Approximate Querying. In: Proceedings of IDEAS 2000, pp. 47–54 (2000b)
5. Intan, R., Mukaidono, M.: Generalization of Fuzzy Rough Sets by Fuzzy Covering Based on Weak Fuzzy Similarity Relation. In: Proceeding of FSKD 2002, pp. 344–348 (2002a)
6. Intan, R., Mukaidono, M.: Degree of Similarity in Fuzzy Partition. In: Pal, N.R., Sugeno, M. (eds.) AFSS 2002. LNCS (LNAI), vol. 2275, pp. 20–26. Springer, Heidelberg (2002b)
7. Intan, R., Mukaidono, M.: Generalized Fuzzy Rough Sets By Conditional Probability Relations. International Journal of Pattern Recognition and Artificial Intelligence 16(7), 865–881 (2002c)
8. Intan, R., Mukaidono, M.: Probability of Fuzzy Event to Probability of Rough Event. In: Proceedings of FSKD 2002, pp. 549–553 (2002d)
9. Intan, R., Mukaidono, M.: A Proposal of Probability of Rough Event based on Probability of Fuzzy Event. In: Alpigini, J.J., Peters, J.F., Skowron, A., Zhong, N. (eds.) RSCTC 2002. LNCS (LNAI), vol. 2475, pp. 357–364. Springer, Heidelberg (2002e)
10. Intan, R., Mukaidono, M.: Hybrid Probabilistic Models of Fuzzy and Rough Events. Journal of Advanced Computational Intelligent 7(3), 323–329 (2003)
11. Inuiguchi, M., Tanino, T.: On Rough Sets under Generalized Equivalence Relations. In: Proceedings of International Workshop on Rough Sets and Granular Computing, pp. 167–171 (2001)
12. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall, New Jersey (1995)
13. Komorowski, J., Pawlak, Z., Polkowski, L., Skowron, A.: Rough Sets: A Tutorial. In: Pal, S.K., Skowron, A. (eds.) Rough Fuzzy Hybridization, pp. 3–98. Springer (1999)
14. Kosko, B.: Fuzziness VS. Probability. International Journal of General Systems 17, 211–240 (1990)
15. Nguyen, H.T.: On Fuzziness and Linguistic Probabilities. Journal of Mathematical Analysis and Applications 61, 658–671 (1977)
16. Pawlak, Z.: Rough sets. International Journal Computation and Information Science 11, 341–356 (1982)
17. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
18. Sugeno, M.: Fuzzy Measures and Fuzzy Integrals, a survey. In: Gupta, Saridis and Gaines, pp. 89–102 (1977)
19. Yao, Y.Y.: Two views of the theory of rough sets in finite universe. International Journal of Approximate Reasoning 15, 291–317 (1996)
20. Zadeh, L.A.: Probability Measures of Fuzzy Events. Journal of Mathematical Analysis and Applications 23, 421–427 (1968)
21. Zadeh, L.A.: Fuzzy Sets and Systems. International Journal of General Systems 17, 129–138 (1990)
22. Zadeh, L.A.: Discussion: Probability Theory and Fuzzy Logic Are Complementary Rather Than Competitive. Technometrics 37(3), 271–276 (1995)