An unified approach to the fusion of imperfect data?

Mihai C. Florea[a], Anne-Laure Jousselme[b], Dominic Grenier[a] and Éloi Bossé[b]

[a] Université Laval
Laboratoire de Radiocommunication et de Traitement de Signal (LRTS)
Département de Génie Électrique, Ste-Foy, Qc, G1K 7P4 Canada

[b] Defence Research Establishment Valcartier, Val-Bélair, Qc, G3J 1X5 Canada

ABSTRACT

For several years, researchers have explored the unification of the theories enabling the fusion of imperfect data and have finally considered two frameworks: the random sets and the conditional events algebra. Traditionally, the information is modeled and fused in one of the known theories: bayesian, fuzzy sets, possibilistic, evidential, or rough sets... Previous work has shown what kind of imperfect data these theories can best deal with. So, depending on the quality of the available information (uncertain, vague, imprecise, ...), one particular theory seems to be the preferred choice for fusion. However, in a typical application, the variety of sources provides different kinds of imperfect data. The classical approach is then to model and fuse the incoming data in a single theory being previously chosen. In this paper, we first introduce the various kinds of imperfect data and then the theories that can be used to cope with the imperfection. We also present the existing relationships between them and detail the most important properties for each theory. We finally propose the random sets theory as a possible framework for unification, and thus show how the individual theories can fit in this framework.

Keywords: Imperfect data, data fusion, random sets.

1. INTRODUCTION

Data fusion is the process of combining data or information to estimate or predict entity states. The resulting imperfect information must be obtained without loosing any information, or with the smallest loss as possible, and most of the time must support a decision making.

Why is data fusion process so difficult? Two aspects are involved: a qualitative and a quantitative one. The problem caused by the qualitative aspect is how to choose best formalism to represent the type of imperfect information and how to characterize this information. There is no formalism to describe all kinds of imperfect information. An imperfect data can be characterized as being imprecise, uncertain, or both. Other imperfections (like vagueness or incompleteness) can be described as a particular case of imprecision and (or) uncertainty (Smets[1]). Once a theory is chosen to describe the data or information, the quantitative aspect is involved. The problem is how to quantify the confidence accorded to a realization (for example: how the belief of an event is set to be 0.8 and not 0.6).

Several theories have been developed to deal with imperfect information. Bayesian theory is known from centuries and deals mainly with uncertainty, but can also deal with imprecision. Rough sets theory (Pawlak[2]) deals with imprecision, when uncertainty is involved but not quantified. Dempster-Shafer's theory of evidence (Shafer,[3] Dempster[4]) is able to deal with information containing imprecision and uncertainty at the same time. On the other hand, the theory of possibility addresses with the incomplete information, which is a combination of imprecision and uncertainty (Zadeh[5]). The theory of fuzzy sets (Zadeh[6]) handles with vague data (vague predicates like "large" or "big" are not well defined in our language). Vague data is a particular case of both imprecise and uncertain data. These theories have been mostly applied to only one type of imperfection. To

Further author information:
Mihai C. Florea: E-mail: mflorea@gel.ulaval.ca, Telephone: 1 418 656 2131 # 4883
Anne-Laure Jousselme: E-mail: annelaur@gel.ulaval.ca, Telephone: 1 418 656 2131 # 4883

cope with all types of imperfection is a much more difficult problem and some potential approaches have been proposed: the theory of random sets and the conditional event algebra (Goodman, Mahler and Nguyen[7]).

Until now, it is not clear how these theories can practically be applied to the fusion of imperfect data. This paper attempts to present and link together the various frameworks dealing with imperfect data as well as better describe the difficulties encountered in the fusion of such information.

Section 2 presents a discussion of the different kinds of imperfect data. Section 3 reviews the theories dealing with the imperfect information. Section 4 discusses the different links between these theories: how, an information that is modeled in one theory, could be formulated in another theory. We will also study the difficulties of applying these transformations and their limits of trust. Finally, in section 5, we introduce the random sets theory and the relationships with the other theories dealing with imperfect information.

2. REPRESENTING IMPERFECT DATA

One can discuss imperfect information in terms of **subjective** or **objective information**. The subjective information is an information which is dependent of the source of information. This means that the source of information has given an interpretation of the data. Usually, humans are the main contributors of subjective information. The subjective information is due to the limitations of language or to the limitations of understanding. An objective information is an information that is not dependent of the source of information.

Let us show an example for these two kinds of information: consider a source of information giving the color of an object. An objective information is the value of the wavelength of the corresponding color. On the other hand, two humans can provide for the same color different characterizations, because the distinction between different shades of green and blue is not well perceived. In this case, the information is depending on the source, so it is a subjective information.

Another aspect of the imperfect information is the degree of **uncertainty** and (or) **imprecision** (Smets[1]). Uncertainty represents our state of knowledge about a piece of information. Imprecision is a characteristic of an information that cannot be expressed by a single value, but by a set of values. According to the situation, uncertainty can present one of the two aspects of subjective or objective information. When the uncertainty characterizes the *chance* then it is an objective information, but when the uncertainty characterizes our *state of truth* then it is an subjective information. Imprecision is a characteristic of an objective information only.

Let us study an example to understand better the difference between uncertainty and imprecision. The information "The detected object is a F-18 plane" is a precise and certain information. This is an perfect information. Unfortunately, in practice, the information is often uncertain and imprecise. The data "The detected object might be a F-18 plane with a degree of confidence of 60%" is an example of an uncertain information. The information is precise but not sure. The confidence that the object detected is a F-18 plane is not 100%. The detected object could be another plane. The information "The detected object is one of the planes {F-18, Boeing 747, F-16, Mig-29}" is an imprecise information. Our state of knowledge is that the detected object is one of the element of the set and we are sure about that. An example of a combination of imprecise and uncertain data is: "The detected object is one of the planes {F-18, Boeing 747, F-16} with a degree of confidence of 60% and one of the planes {Mig-29, F-15, F-22 } with a degree of confidence of 40%".

Usually, the information provided by a sensor is both imprecise and uncertain, and a relationship exists between the two kinds of imperfection: the more an information is precise, the more its belief reduces. If we can access the raw data, then usually we can manage the sensor to get the kind of information we need (more precise or more certain).

Another kind of imperfect data is the **vague information**. This kind of data is due specially to the limitation of the vocabulary and is in the most of the cases a subjective data. What is a vague data? It is an information which describes a class of objects, but the limits of this class are not well known. The information "Tom died *young*" is an information sure but vague because we do not know how old was Tom when he died, neither what are the limits of the "*young*" classification. Does *young* mean more $age = 15$, or $age = 25$, or more $age = 50$? What are the limits of "*young*" and how can we grant our confidence to each value? This vague data can be considered as a particular case of imprecise and uncertain information and is processed in the theory of fuzzy sets.

Another kind of data is **incomplete information**. This case is also a combination of imprecise and uncertain information. An incomplete data is represented by the upper limit of our degree of confidence (called also the possibility) of an event. This means we do not know the probability of an event but we know its possibility of realization. The difference can be illustrated by the example used by Zadeh[5] : "Hans eats X eggs" with $X \in \{1, 2, 3, 4, ...\}$. The possibility that Hans eats 3 eggs is 1, but the probability that Hans eats 3 eggs is only 0.1. Both information are different. The degree of possibility is only an incomplete information which is less expressive than the probability.

The last type of imperfect information is the **inconsistent information**. This aspect is not connected to neither the imprecision or uncertainty, but is connected to the prior information. We cannot say about a single piece of information (an information by itself) that it is inconsistent. It might be compared to another piece of information. Two pieces of information can be inconsistent even though it were declared to be precise and certain by two different sources. For example, if two sensors have detected the color of an object to be green, respectively red, even if the two sensors are precise and certain, we can conclude that the pieces of information provided are inconsistent. If another piece of information is received to confirm one or another piece of the inconsistent information, then we can eliminate the wrong one. Otherwise, we can only create an imprecise data from the two inconsistent information. Using the last example from the two incoherent information we can create the following information "The color of the detected object is green or red".

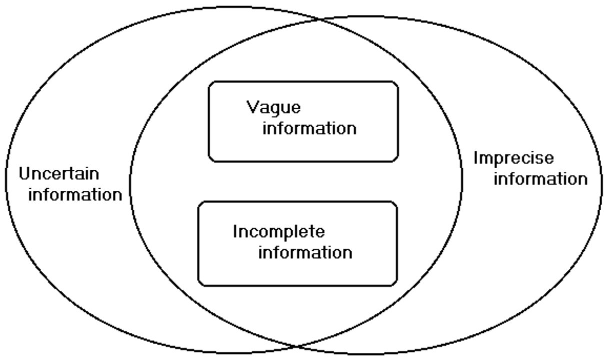

Figure 1: Different kinds of imperfect information.

The figure 1 summarizes the different kinds of imperfect information and the relationships existing between them.

3. TOOLS FOR DEALING WITH THE IMPERFECT INFORMATION

In the last 30 years, the growing knowledge about how the information can be mathematically represented, permitted the development of new theories dealing with the imperfect information. Thus, moreover the theory of probability already known, several theories were born: fuzzy sets theory (1965), theory of evidence (1976), theory of possibility (1978) and rough sets theory (in the '80s). These new theories were developed to characterize in a more natural way all the aspects of data's imperfections. It is now clear that a lot of cases of imperfect information can be modeled by several of the theories listed above. The goal of these theories is to imitate as well as possible (in the best natural way) the human reasoning.

Before describing the objective of each theory in particular, we will present some considerations of vocabulary and notations.

Let us denote by Θ the frame of discernment, $\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$. This is the set of all possible realizations. The power set 2^Θ is the set of all subsets of Θ. This power set contains also the empty set $\varnothing$ and the frame of discernment Θ. Thereafter, we will denote par θ an element of Θ ($\theta \in \Theta$) and by A an element of the power set ($A \in 2^\Theta$ or $A \subseteq \Theta$). $\cup$ means the union of two or more events, while $\cap$ means the intersection of two or more events. In other words, $\theta_1 \cup \theta_2$ means "θ_1 *or* θ_2", while $\theta_1 \cap \theta_2$ means "θ_1 *and* θ_2".

3.1. Probability theory

The probability theory was for a long time used to deal with almost all kinds of imperfect information, because it was the only existing theory. The measure of **probability** expresses the degree of confidence that someone grants to the occurrence of a realization of an event (the union of all the possible events are forming the frame of discernment Θ). The probability measure, P, satisfies three axioms:

$$P[\varnothing] = 0 \tag{1}$$

$$0 \leq P[A] \leq 1, \forall A \in 2^{\Theta} \tag{2}$$

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] \tag{3}$$

The first axiom corresponds to the assumption of a close world, which mean that all the possible realizations of an unknown event are described in the frame of discernment. If $A \cap B = \varnothing$ then the 3rd axiom become:

$$P[A \cup B] = P[A] + P[B] \tag{4}$$

If the realization of two events ere independent, then:

$$P[A \cap B] = P[A]P[B] \tag{5}$$

The knowledge about the realization of the event A, can change the knowledge about the realization of another event B by (this process is also known by Bayes law of conditioning):

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \tag{6}$$

Ignorance is modeled by granting the same probability to all possible realizations of an event:

$$P[\theta_1] = P[\theta_2] = ... = P[\theta_n] = \frac{1}{n} = \frac{1}{\mathrm{card}(\Theta)} \tag{7}$$

The fusion of information coming from different sensors is expressed by the formula of total probability:

$$P[A] = \sum_{j=1}^{m} P[A|source_j]P[source_j] \tag{8}$$

where $P[A|source_j]$ represents the information offered by the source j. One difficulty with the probability theory is to consider the information sources equiprobable. This happens when there exists no other information about the reliability of sources.

At the beginning, this theory was developed to deal with random experiments involving combinatory logic theory or involving the frequentist limits. The probability theory is a very good tool for solving this kind of uncertainty problems. This theory can also deal with imprecision, but the probability of an imprecise event is strongly dependent of the probabilities of precise events (see Eq. 4). In a lot of cases, the prior information is not available and the user does not have all the data to solve the problem. Moreover, imperfect information, especially the imprecise one, is hardly modeled with the probability theory. For this kind of imperfect information, other theories were proposed.

3.2. Theory of evidence

The Dempster-Shafer's theory of evidence (Shafer[3]) was developed as an alternative to the probability theory to fulfill the need of dealing with both imprecision and uncertainty. One of the specificities of this theory is the frame of work that is no more the frame of discernment, but the power set 2^{Θ}. For each subset of the frame of discernment, we consider a **basic probability assignment**, m, that must satisfies two conditions:

$$m(\varnothing) = 0 \tag{9}$$

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{10}$$

where $m(A)$ represents the confidence we grant to the realization of A and only of A.

We can easily see that the evidential theory is reduced to the probability theory if the imprecision is suppressed: i.e. the basic probability assignment is defined only for the singletons of Θ.

The **belief** and the **plausibility** are two other functions of the evidencial theory:

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(B) \tag{11}$$

$$\mathrm{Pl}(A) = \sum_{A \cap B \neq \varnothing} m(B) \tag{12}$$

These two functions express the total belief granted to the event A - $\mathrm{Bel}(A)$, respectively the limit of the belief of the event A - $\mathrm{Pl}(A)$).

Ignorance is modeled by granting the total belief to the frame of discernment $m(\Theta) = 1$ and $m(A) = 0$ $\forall A \subset \Theta$.

The fusion of two basic probability assignments (m_1 and m_2) is realized using the Dempster's rule of combination (Dempster[4]):

$$m(A) = \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \varnothing} m_1(A_i) m_2(B_j)} \tag{13}$$

where m is the basic probability assignment obtained after the fusion.

3.3. Rough sets

The rough sets theory (Pawlak[2]) is dealing with the aspect of imprecision. The basic concept of the rough sets theory is to replace an uncertain imprecise information by two imprecise but certain information: the **lower** and **upper approximations**. The combination of imprecise information is realized by applying the set theory to the approximations. The beauty of the rough sets theory is that there is no need to quantify the information's uncertainty. This is an advantage because usually it is very difficult to quantify the degree of confidence granted to an information, and it is a disadvantage because there is no difference between two pieces of information to which we would have granted different degrees of confidences.

In fact, the lower and upper approximations are combinations between the uncertain information and the prior knowledge about the frame of discernment. A prior knowledge is defined as a subset of 2^Θ and describes a classification of the realizations from the frame of discernment. Usually, this knowledge is a partition of the frame of discernment, which is a particular subset of the power set.

Let $R \subseteq 2^\Theta$ be a prior knowledge describing the frame of discernment Θ. The lower and upper approximations of an imprecise information, in agreement with the knowledge R are:

$$\underline{R}A = \{[\theta]_R | [\theta]_R \subseteq A, \theta \in \Theta\} \tag{14}$$

$$\overline{R}A = \{[\theta]_R | [\theta]_R \cap A \neq \emptyset, \theta \in \Theta\} \tag{15}$$

where $[\theta]_R$ represents the subset of 2^Θ, also a subset of R, that contains θ.

Combining two uncertain and imprecise information A and B, described by two pieces of certain information $(\underline{R}A, \overline{R}A)$, respectively $(\underline{R}B, \overline{R}B)$, resides in finding the lower and upper approximations of $A \cup B$ and $A \cap B$:

$$\underline{R}(A \cap B) = \underline{R}A \cap \underline{R}B \tag{16}$$

$$\overline{R}(A \cap B) \subseteq \overline{R}A \cap \overline{R}B \tag{17}$$

$$\underline{R}(A \cup B) \supseteq \underline{R}A \cup \underline{R}B \tag{18}$$

$$\overline{R}(A \cup B) = \overline{R}A \cup \overline{R}B \tag{19}$$

A membership function of an imprecise information $A \subseteq \Theta$ was defined by Pawlak[8]:

$$\mu_A(\theta) = \begin{cases} 1 & if \;\; \theta \in \underline{R}A \\ 0.5 & if \;\; \theta \in [\overline{R}A - \underline{R}A] \\ 0 & if \;\; \theta \in -\overline{R}A \end{cases} \tag{20}$$

3.4. Fuzzy sets

Fuzzy sets theory is based on the concept of uncertain membership to a set. In the general case, a subset $A \subseteq \Theta$ can be represented by a membership function μ_A:

$$\forall \theta \in \Theta \quad \mu_A(\theta) = \begin{cases} 1 & \text{if } \theta \in A \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

Let us note $\underline{A}$ an uncertain subset of Θ, called also a fuzzy set. Then, its membership function is $\mu_{\underline{A}}(\theta) \in [0, 1]$ $\forall \theta \in \Theta$, where $\mu_{\underline{A}}(\theta)$ describes the degree of membership of θ to $\underline{A}$. The theory of fuzzy sets is dealing especially with vague, ill defined, or ambiguous data. This concept of fuzzy sets - fuzzy membership is the most natural way for describing this kind of imperfect information. In this theory, θ does not play the role of the imperfect information, but the role of precise information. The role of imperfect information is played by the fuzzy set $\underline{A}$, which is considered as ill defined, and θ does not certainly belong to $\underline{A}$ because its limits are not well defined.

Combining several fuzzy information is combining the fuzzy sets that best describe the information. For example if two sensors describe a detected object as red and big (both qualifiers are vague data), then combining these information means to combine the two fuzzy sets that each information generated. There exists several ways to combine two fuzzy sets, but we are describing here only a few.

The intersection of two fuzzy sets $\underline{A}$ and $\underline{B}$:

$$\forall \theta \in \Theta \quad \mu_{\underline{A} \cap \underline{B}}(\theta) = \min\{\mu_{\underline{A}}(\theta), \mu_{\underline{B}}(\theta)\} \tag{22}$$

The union of two fuzzy sets $\underline{A}$ and $\underline{B}$:

$$\forall \theta \in \Theta \quad \mu_{\underline{A} \cup \underline{B}}(\theta) = \max\{\mu_{\underline{A}}(\theta), \mu_{\underline{B}}(\theta)\} \tag{23}$$

The algebraic product of two fuzzy sets $\underline{A}$ and $\underline{B}$:

$$\forall \theta \in \Theta \quad \mu_{\underline{A}.\underline{B}}(\theta) = \mu_{\underline{A}}(\theta)\mu_{\underline{B}}(\theta) \tag{24}$$

The algebraic sum of two fuzzy sets $\underline{A}$ and $\underline{B}$:

$$\forall \theta \in \Theta \quad \mu_{\underline{A} \hat{+} \underline{B}}(\theta) = \mu_{\underline{A}}(\theta) + \mu_{\underline{B}}(\theta) - \mu_{\underline{A}}(\theta)\mu_{\underline{B}}(\theta) \tag{25}$$

The combinations obtained from the Luckasiewicz implication:

$$\forall \theta \in \Theta \qquad \mu_{\underline{A} \cap \underline{B}}(\theta) = \max\{0, \mu_{\underline{A}}(\theta) + \mu_{\underline{B}}(\theta) - 1\} \tag{26}$$

$$\mu_{\underline{A} \cup \underline{B}}(\theta) = \min\{1, \mu_{\underline{A}}(\theta) + \mu_{\underline{B}}(\theta)\} \tag{27}$$

These combination methods are different and are applied to different situations according to the relationships between information (correlated information, incoherent information), and the result we want to obtain (conjunctive or disjunctive combinations).

3.5. Possibility theory

The possibility theory is a tool able to deal with both imprecision and uncertainty. For every $\theta \in \Theta$ is defined a possibility measure, which represents a limit of our degree of confidence given to it. The possibility measure, $\pi(\theta)$, takes values in the [0,1] interval; $\pi(\theta) = 1$ means that the event is possible, but does not mean that it is also certain.

The possibility theory can treat the imperfect data in two different ways:

- when the measures of possibility are available we are talking about the numerical aspect of the possibilities;

- when the measures of possibility are not available, then the possibility theory deals only with the relations existing between the possibilities (a minimum information must be known on the order between the possibilities).

For every imprecise event $A \subseteq \Theta$ there exists two measures to qualify the confidence accorded to it - the possibility degree, $\Pi(A)$, and the necessity degree, $N(A)$:

$$\Pi(A) = \sup_{\theta \in A}\{\pi(\theta)\} \tag{28}$$

$$N(A) = 1 - \Pi(\overline{A}) \tag{29}$$

Several methods for combining the information were proposed, but some of them are not even distributive, a propriety that is essential in data fusion. Some of the conjunctive ($\pi_a^{\cap}$, $\pi_b^{\cap}$, $\pi_c^{\cap}$) and disjunctive ($\pi_a^{\cup}$, $\pi_b^{\cup}$, $\pi_c^{\cup}$) combinations are presented below:

$$\pi_a^{\cap}(\theta) = \frac{\min\{\pi_1(\theta), \pi_2(\theta)\}}{h(\pi_1, \pi_2)} \qquad \text{with} \qquad h(\pi_1, \pi_2) = \sup_{\theta \in \Theta} \min\{\pi_1(\theta), \pi_2(\theta)\} \tag{30}$$

$$\pi_b^{\cap}(\theta) = \min\{\pi_1(\theta), \pi_2(\theta)\} + 1 - h(\pi_1, \pi_2) \tag{31}$$

$$\pi_c^{\cap}(\theta) = \max\{0, \pi_1(\theta) + \pi_2(\theta) - 1\} \tag{32}$$

$$\pi_a^{\cup}(\theta) = \max\{\pi_1(\theta), \pi_2(\theta)\} \tag{33}$$

$$\pi_b^{\cup}(\theta) = \pi_1(\theta) + \pi_2(\theta) - \pi_1(\theta) \times \pi_2(\theta) \tag{34}$$

$$\pi_c^{\cup}(\theta) = \min\{\pi_1(\theta) + \pi_2(\theta), 1\} \tag{35}$$

Each of these combinations can be used in particular conditions and also with specific purposes: one can eliminate the less possible realizations (Eq. 32), the conjunctive ones are applying when the two sensors provide reliable information, since the disjunctive ones are applied when one of the sensors provides an erroneous information.

Dubois & Prade[9] proposed a new method doing an adaptive combination (a mixture between a conjunctive and a disjunctive combination) which can apply in almost all the situations:

$$\pi_{AD}(\theta) = \max\left(\frac{\pi^{\cap}(\theta)}{h(\pi_1, \pi_2)}, \min(1 - h(\pi_1, \pi_2), \pi^{\cup}(\theta))\right) \tag{36}$$

where $h(\pi_1, \pi_2)$ is given in Eq. 30.

4. LINKS BETWEEN THEORIES

Several theories were developed to deal with imperfect data, but none of these theories is able to deal with all kinds of imperfect information. Transformations between theories need to be considered. When passing from a theory to another, some information is lost and imprecision is increased. Below is a survey of different links between the various theories.

4.1. Probabilities vs. evidence theory

Probability theory may be seen as a restriction of the evidence theory since if the basic probability assignment is defined only for singletons ($\theta \in \Theta$), an equivalence can be established between the two theories:

$$P[X = \theta] = m(\theta) = \mathrm{Bel}(\theta) = \mathrm{Pl}(\theta) \quad \forall\ \theta \in \Theta \tag{37}$$

The evidence theory is more expressive than the probability theory, and passing from one to another is possible only with a loss on information. The pignistic transformation (Smets[10]) is one of the most known transformation from the beliefs to the probabilities, and it can be used only in this direction:

$$P[X = \theta] = \sum_{\theta \in A,\ A \subseteq \Theta} \frac{m(A)}{\mathrm{card}(A)} \qquad \forall \theta \in \Theta \tag{38}$$

To several representations of information in the evidential theory corresponds only one representation in the bayesian theory. The mapping between all the representations supported by the two theories is not bijective. A transformation from a belief function to a probability will always introduce a loss of information. From probability to belief, any transformation will increase the imperfection of the information because of an arbitrary choice that have to be made.

4.2. Evidence theory vs. rough sets theory

A very useful link between these two theories was realized by Skowron[11] and Skowron & Grzymala-Busse.[12] They showed that every problem modelized by the rough sets theory has a corresponding modelization in the theory of evidence. They create the uncertainty measures in the evidence theory (the belief and the plausibility) using the prior knowledge $R \subseteq 2^{\Theta}$:

$$\mathrm{Bel}(A) = \frac{\mathrm{card}(\underline{R}A)}{\mathrm{card}(\Theta)} \tag{39}$$

$$\mathrm{Pl}(A) = \frac{\mathrm{card}(\overline{R}A)}{\mathrm{card}(\Theta)} \tag{40}$$

4.3. Fuzzy sets vs. possibility

Zadeh[5] has developed the theory of possibilities from the theory of fuzzy sets. He proved that the two theories are strongly linked. For Zadeh, the membership function of a fuzzy set expresses in the same time the distribution of possibility:

$$\mu_{\underline{A}}(\theta) = \pi_X(\theta) \tag{41}$$

Zadeh's interpretation is that $\pi_X(\theta)$ measures the possibility that $X = \theta$ on a basis of the proposition "X is $\underline{A}$".

Dubois & Prade[13] propose another interpretation of the membership function in term of possibility measure, a conditional possibility:

$$\mu_{\underline{A}}(\theta) = \Pi(\underline{A}|\theta) \tag{42}$$

These different valid interpretations of the membership function in terms of possibility measure can show that the human reasoning is still difficult to imitate in mathematical equations and the models proposed so far are not perfect.

4.4. Possibility vs. probability

After doing a revision of different transformations between the probability and the possibility theories, Klir & Parviz[14] have concluded that on some aspects the probabilities are better, while on other aspects the possibilities are better and finally that the theories are not comparable.

This same conclusion was expressed also by Zadeh[5] after he studied the example: "Hans eat X eggs" with $X \in \Theta = \{1, 2, 3, 4, ...\}$:

θ	1	2	3	4	5	6	7	8
$\pi(\theta)$	1	1	1	1	0.8	0.6	0.4	0.2
$p(\theta)$	0.1	0.8	0.1	0	0	0	0	0

Table 1: Difference between probability and possibility measures

The measures of probability and the possibility presented in Table 1 are totally different. We cannot say that one measure can replace another, or that one measure can be transformed into another, because it is obvious that both measures do not express the same information.

4.5. Possibility vs. evidence theory

Comparing the couples of measures (possibility, necessity) and (plausibility, belief) Dubois & Prade[15, 16] concluded that the possibility theory is a particular case of theory of evidence (when the basic probability assignment is defined only for nested sets). The measures are expressing in fact the same thing, but the possibility theory can deal with imperfect data even if we do not know the possibility measure.

On the other hand, Smets[17] and Sudkamp[18] disagree with the conclusion of Dubois & Prade. They consider that although the mathematics are the same, the two sets of measures are applied to different kinds of information.

4.6. Fuzzy sets vs. rough sets

Pawlak[8] proved that the membership function of a fuzzy set could not be compared with the membership function of a rough set (Eq. 20) because the two functions do not behave in the same manner with the operations of union and intersection of the sets.

Dubois & Prade[19] criticize the manner Pawlak compares both theories. They propose not to consider them as rival theories but to combine them into two different ways to express, in a more naturally way, a complex imperfect information: fuzzy rough sets and rough fuzzy sets.

The various transformations from one theory to another have some important consequences: the loss of information or the increase of imperfection (uncertainty or imprecision). We must use them with care.

5. RANDOM SETS AND LINKS WITH THE OTHER THEORIES

The concept of random sets was developed for geometry applications by Kendall[20] in the 70's. Some authors have viewed the random sets as a method of studying imperfect information and have made links between random sets and other existing theories: Nguyen,[21] with the evidence theory, and Orlov,[22] with the fuzzy sets. Other researchers have used the random sets formalism as an unification method for dealing with imperfect information : Mahler,[23] Mori,[24] Kreinovich.[25]

What is a random set ? Let $[X, \mathcal{A}, P]$ be a finite probability space and let P be a probability function defined by:

$$\begin{aligned} P \quad : \quad & X \longrightarrow [0,1] \\ & x \longmapsto P[x] \end{aligned} \tag{43}$$

If M is a random variable defined by the mapping:

$$\begin{aligned} M \quad : \quad & X \longrightarrow 2^{\Theta} \\ & x \longmapsto M[x] = A \end{aligned} \tag{44}$$

then M constitutes a **random set**.

5.1. Link with the Dempster-Shafer's theory of evidence

If we define a basic probability assignment m on 2^{Θ} we can state that:

$$m(A) = P[M^{-1}(A)] \tag{45}$$

and it can be said that M is a random set with m as probability distribution.

The link between the random sets theory and the Dempster-Shafer theory has been established by (Nguyen,[21] Nguyen & Wang[26]):

$$\text{Bel}(A) = P[M \subseteq A] \tag{46}$$

$$\text{Pl}(A) = P[M \cap A \neq \varnothing] \tag{47}$$

Nevertheless, a condition must be satisfied to have an equivalence between the two concepts:

$$P[M = \varnothing] = 0 \tag{48}$$

So, an information represented in the theory of evidence can be expressed by a set:

$$\begin{cases} \{A, m(A)\} \ \forall \ A \subseteq \Theta \\ \sum_{A \subseteq \Theta} m(A) = 1 \end{cases} \tag{49}$$

which is also a representation for the random set formalism.

Even more, if we are studying the Dempster's rule of combination (Eq. 13), it can be express in term of probability of a random set:

$$P[M = A] = \frac{P[M_1 \cap M_2 = A]}{P[M_1 \cap M_2 = \varnothing]} \tag{50}$$

5.2. Link with the fuzzy sets theory

For better seeing the link between the fuzzy sets and the random sets, we must introduce another concept in the fuzzy sets theory - the α level cut of an fuzzy set $\underline{A}$, noted $\underline{A}_\alpha$ (Dubois & Prade[16]):

$$\underline{A}_\alpha = \{\theta \in \Theta | \mu_{\underline{A}}(\theta) \geq \alpha\} \quad \forall \alpha \in [0,1] \tag{51}$$

which is a crisp set.

Then, for a set of values $1 = \alpha_1 > \alpha_2 > .... > \alpha_n > 0$ we obtain the α level cuts for a fuzzy set $\underline{A}_{\alpha 1} \subseteq \underline{A}_{\alpha 2} \subseteq ... \subseteq \underline{A}_{\alpha n} = \Theta$ and then the fuzzy set can be represented as a set of couples:

$$\begin{cases} \{\underline{A}_{\alpha i}, \beta_i\} \text{ with } \beta_i = \alpha_i - \alpha_{i+1} \\ \sum \beta_i = 1 \end{cases} \tag{52}$$

This is also a representation of the same information in the random set formalism.

So far, the most common imperfect information (imprecise/uncertain and vague data) can be modelized in a natural way by the theory of random sets.

6. CONCLUSION

This paper presented a summary of the different theories dealing with the imperfect information, their limitations, and the relationship between them. In the final section, the random sets formalism is proposed as an unification framework to the data fusion problem. Even if random sets formalism is relatively new and not very well known, the links with the Dempster-Shafer's theory of evidence and with the fuzzy sets theory make it a powerful tool for data fusion. Future work will concentrate on how this new theory can be applied to the real world situations.

ACKNOWLEDGMENTS

We thank the Defence Research Establishment of Valcartier for the financial and technical support during the course of this work.

REFERENCES

1. P. Smets, "Imperfect Information : Imprecision - Uncertainty," in *Uncertainty Management in Information Systems. From Needs to Solutions*, A. Motro and P. Smets, eds., pp. 225–254, Kluwer Academic Publishers, 1997.
2. Z. Pawlak, *Rough Sets - Theorical Aspects of Reasoning about Data*, Kluwer Academic Publisher, 1991.
3. G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.
4. A. Dempster, "Upper and Lower Probablilities Induced by Multivalued Mapping," *Ann. Math. Statist.* **38**, pp. 325–339, 1967.
5. L. A. Zadeh, "Fuzzy Sets as a Basis for a Theory of Possibility," *Fuzzy Sets and Systems* **1**, pp. 3–28, 1978.
6. L. A. Zadeh, "Fuzzy Sets," *Information and control* **8**, pp. 338–353, 1965.
7. I. Goodman, R. P. S. Mahler, and H. Nguyen, *Mathematics of data fusion*, Kluwer Academic Publishers, Dordrecht, 1997.
8. Z. Pawlak, "Rough Sets and Fuzzy Sets," *Fuzzy Sets and Systems* **17**, pp. 99–102, 1985.
9. D. Dubois and H. Prade, "La fusion d'informations imprécises," *Traitement du Signal* **11**(6), pp. 447–458, 1994.
10. P. Smets, "Belief functions versus probability functions," in *Uncertainty and intelligent systems*, pp. 17–24, Springer Verlag, 1988.
11. A. Skowron, "The Relationship Between the Rough Sets and the Evidence Theory," *Bulletin of the Polish Academy of Science* **37**(1-2), pp. 87–90, 1989.

12. A. Skowron and J. Grzymala-Busse, "From Rough Sets Theory to Evidence Theory," in *Advances in the Dempster-Shafer Theory of Evidence*, pp. 193–235, John Wiley and Sons, NY, 1994.
13. D. Dubois and H. Prade, "Fuzzy Sets and Probability: Misunderstandings, Bridges and Gaps," in *Proceedings of the Second IEEE International Conference on Fuzzy Systems, IEEE* , pp. 1059–1068, 1993.
14. G. J. Klir and B. Parviz, "Probability-Possibility Transformations : A Comparison," *International Journal of General Systems* **21**, pp. 291–310, 1992.
15. D. Dubois and H. Prade, "On Several Representations of an Uncertain Body of Evidence," in *Fuzzy Information and Decision Processes*, pp. 167–181, North Holland, 1982.
16. D. Dubois and H. Prade, *Fundamentals of Fuzzyy Sets*, Kluwer Academic Publishers, 2000.
17. P. Smets, "The Transferable Belief Model and Possibility Theory," *Proc. NAFIPS* , pp. 215–218, 1990.
18. T. Sudkamp, "The Semantics of Plausibility and Possibility," *International Journal of General Systems* **21**, pp. 273–289, 1992.
19. D. Dubois and H. Prade, "Rough Fuzzy Sets and Fuzzy Rough Sets," *International Journal of General Systems* **17**, pp. 191–209, 1990.
20. D. G. Kendall, "Foundation of a theory of random sets," in *Stochastic Geometry*, pp. 322–376, John Wiley & Sons, 1974.
21. H. T. Nguyen, "On Random Sets and Belief Functions," *Journal of Mathematical Analysis and Applications* **65**, pp. 531–542, 1978.
22. A. I. Orlov, "Fuzzy and Random Sets," *Prikladnoi Mnogomerni Statisticheskii Analys* , pp. 262–280, 1978.
23. R. P. Mahler, "Random Sets in Information Fusion. An Overview," in *Random Sets : Theory and Applications*, pp. 129–164, Springer, 1997.
24. S. Mori, "Random Sets in Data Fusion. Multi-Objet State-Estimation as a Foundation of Data Fusion Theory," in *Random Sets : Theory and Applications*, pp. 185–207, Springer, 1997.
25. V. Kreinovich, "Random sets unify, explain, and aid known uncertainty methods in expert systems," in *Random Sets : Theory and Applications*, pp. 321–345, Springer, 1997.
26. H. Nguyen and T. Wang, "Belief Functions and Random Sets," in *Random Sets : Theory and Applications*, pp. 243–255, Springer, 1997.