2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

The Properties and Information Measures for Information Sets

Manish Agarwal, Madasu Hanmandlu , and Kanad K. Biswas

Abstract—**Information sets expand the scope of the existing uncertainty theories by representing uncertainty in the information sources. An information set takes into account both the information source values as well as the membership grades. In this study, we investigate in detail the properties of the information sets. The information measures for the information sets are also developed to measure the information content in an information set.**

I. Introduction

The fuzzy sets theory helps to model the vague concepts associated with the attributes of objects or classes using the membership functions (MFs). The notion of membership functions has been introduced to deal with the foregoing form of uncertainty of classification [1]. They range between 0 and 1 reflecting an assessor's view about the partial degree of membership in an imprecise class. The MF caters to the uncertainty of classification which arises due to the imprecise boundaries of the classes. Since, the appearance of fuzzy set theory, it has been applied in a number of problems [7-11].

At the same time many researchers have cited the problem of interpretation in fuzzy membership function [12-14]. For instance, suppose an observer in Mali conveys to a receiver in Alaska that temperature in Mali is hot, it might be the case that the receiver in Alaska may not get the right reference with the value hot. The same stands true with the membership functions as well. Also, the fuzzy set theory treat the information value and the respective membership function separately. The membership functions only give the belongingness of an attribute to the concept but do not make use of the attributes. The membership functions express the degree of association of each attribute to the respective concept, which is qualitative information.

The concept of information sets has been introduced in [2] to address these gaps by taking into account the attribute values along with the membership grades. This minimizes the subjectivity in the representation of uncertain concepts. The information set gives a uniquely interpretable mapping between information source and its evaluation in terms of of entropy (information) as evaluated by an agent, through a flexible parameter based framework. The capabilities of the agent in representing uncertainty go far beyond those of the membership functions [2].

In this work, we present the properties of the information sets. The information measures for the proposed information sets are also developed to measure the information content in the given information sets. Next, we interpret the information measures with the aid of three examples to illustrate the concepts.

The rest of the paper is organized as follows. Section 2 discusses the notations. Section 3 gives an overview of fuzzy and information sets. Section 4 investigates the properties of the information sets. In Section 5, we present the information measures for the information sets. The proposed information measures are computed for three different combinations of fuzzy sets. Section 6 concludes the paper.

II. Notations

We consider a tabular representation of an information system, as shown in Fig. 1. Each row corresponds to an object and each column corresponds to an attribute. A cell (ij) of this representation gives the information value that ith object takes for jth attribute. All such values in a column represent the possible values for an attribute, and constitute an information source. We use the following notations throughout the study to define an information table:

$$T = (U, At, \{I_X | X \in At\}) \tag{1}$$

where, $U = \{x_1, \dots, x_n\}$ is a finite set of objects,

At is a finite set of attributes

I_X is an information source of $X \in At$, comprising of the values that objects

$\{x_i\} \in U$ take for X.

When $\{x_i\}, i = 1, \dots, n$ are clustered, it will lead to a certain number of fuzzy sets, as shown in Fig. 2. We denote these fuzzy sets by $F_X^k = \{I_X^k(x_i^k), \mu_X^k(x_i^k)\}, k = 1, \dots, z$, $x_i \epsilon U$, $\mu_X^k(x_i^k) \epsilon [0,1]$.

Here, x_i^k refers to those elements x_i from U, which belong to fuzzy set k; $I_X^k(x_i^k)$ refers to the information source for x_i^k; F_X^k stands for fuzzy set k; z is the number of such fuzzy sets F_X^k, and the set $\{\mu_X^k(x_i^k)\}$ represents the MF of $\{x_i^k\}$ in fuzzy set k, generated from the information source I_X^k. An information source I_X^k is a set of values that attribute X takes elements x_i^k, belonging to fuzzy set k. For example, the components of multimedia elements, viz., speech, image, text, video constitute such information sources.

In the context of an image, the components of each of the color models RGB, HSV constitute the information sources. For an object x_i^k in fuzzy set k, $I_X^k(x_i^k)$ refers to the information source value I_X^k for x_i^k. In the light of the above notations, we can write

$$I_X^k = \{I_X^k(x_i^k)\}, i = 1, \dots, n \tag{2}$$

An information source may also represent values for an attribute that may be varying in space or time. All such values of an attribute changing over time or space would constitute an information source. For instance, the changes in a parameter values due to face change, as a person gets

M. Agarwal is with Department of Computer Science and Engineering, IIT Delhi (corresponding author; e-mail: magarwal@ cse.iitd.ernet.in).
M. Hanmandlu, is with Department of Electrical Engineering, IIT Delhi
K. K. Biswas is with Computer Science and Engg. Department, IIT Delhi.

older, generate an information source. A data element, $x_i \epsilon U$, can be represented as a union of such fuzzy sets, F_X^k.

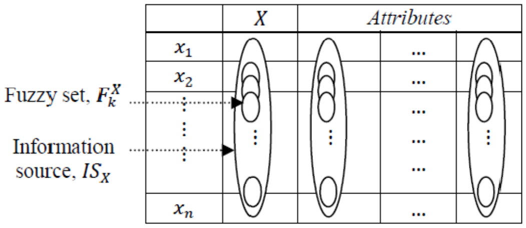

Fig. 1: Information Table with Information Source

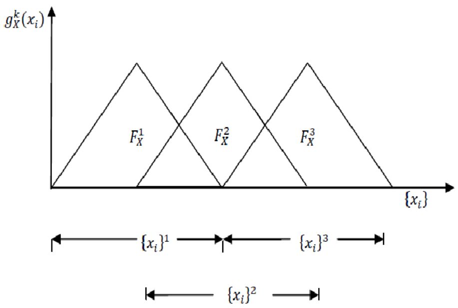

Fig. 2: The Corresponding Fuzzy Sets for an Information Source

III. FROM FUZZY SETS TO INFORMATION SETS

The fuzzy set theory is the generalization of classical set theory to deal with uncertainty [3]. A fuzzy set may be viewed as a class with non-sharp boundaries [4], addressing the issue of granulation in terms of fuzzy sets with a soft transition between the classes in the case of fuzzy sets.

Let $U = \{x_1, \dots, x_n\}$ is a finite and non-empty set of objects, and $\mathcal{F}(\mathrm{U})$ denotes the set of all fuzzy subsets from U. For all the values that an attribute X takes for the objects, $\{\mathrm{x}_1, \dots, \mathrm{x}_\mathrm{n}\}$, a fuzzy set k is defined by a membership function μ_X^k from the values of X, i.e. $\mu_X^k \to [0,1]$. A fuzzy set k of U may also be thought of as a mathematical representation of a ''vague'' concept described linguistically. The fuzzy set theory leads to the generalized notions of sets and members of sets, compared with classical sets [5]. This observation leads to the view that the theory of fuzzy sets emerges from a deviation of classical set theory. The core and support are qualitative representations of a fuzzy set. They can be used as approximations of a fuzzy set [6]. They are defined as

$$core(\mu_X^k) = \{\mu_X^k(x_i^k) = 1\}$$
$$support(\mu_X^k) = \{\mu_X^k(x_i^k) > 0\} \tag{3}$$

A fuzzy set models the ill-definition of a boundary of an imprecise concept, described linguistically, through MFs. A rough set characterizes an otherwise precise concept by its lower and upper approximations in terms of equivalence classes, because of incomplete information. The core and support of fuzzy set theory may be referred to as the lower and upper approximations of a set, in the rough set theory. For example, in the fuzzy set theory, pleasant weather is represented by different membership values in the classes "hot weather" and "cold weather". On the other hand, these are visualized as the lower and upper bounds for the pleasant weather in a rough set.

The information set [2] is constructed by taking into account the information source in a set. The membership function is assigned by an agent to the set. For situations, where an agent does not have an adequate knowledge to assess the information source precisely, the information set should come to his rescue. For example, let an agent assesses the distance between two cities as approximately 150 Km with a membership grade of 0.65 in the fuzzy set "near". In the original fuzzy set theory, the information is communicated in terms of the membership grade only, whereas the proposed information set represents the information source (150 Km), conditioned by the membership function (0.65) by the agent. The information associated with the information set in this case is 97.5. The concept of information arises from representing the uncertainty in an information source through the notion of information gain.

Definition 3.1 [2]: The information gain corresponding to x_i is defined as

$$g_X^k(x_i) = e^{-\left(a(I_X(x_i))^3 + b(I_X(x_i))^2 + c(I_X(x_i)) + d\right)^\alpha} \tag{4}$$

where $g_X^k(x_i)$ is the information gain due to x_i in k^{th} fuzzy set, and a, b, c, d and α are real valued parameters.

Definition 3.2 [2]: The information set S_X^k corresponding to information source I_X and fuzzy set k is given as

$$S_X^k = \{I_X^k(x_i^k) g_X^k(x_i^k)\}, \forall x_i \in k, U \tag{5}$$

An information set corresponding to a fuzzy set, comprises of the entropy values for various elements of the fuzzy set. In the domain of information set theory, the entropy for an element of a fuzzy set is defined as the product of information gain and information source. For a given attribute X and a fuzzy set k, the entropy corresponding to object x_i^k forms an element of information set S_X^k, and is expressed as

$$S_X^k(x_i^k) = \left(I_X^k(x_i^k) g_X^k(x_i^k)\right) \tag{6}$$

The information gain here could be replaced with the membership function in case a well defined distribution is not available.

Definition 3.3 [2]: The information H_X^k associated with S_X^k is given as

$$H_X^k = \frac{1}{|F_X^k|} \sum_i I_X^k(x_i^k) g_X^k(x_i^k), \forall x_i \in k, U \tag{7}$$

where, the term $|F_X^k|$ refers to the cardinality (the count of the elements) of F_X^k.

Thus the original information source I_X gets translated into the information H_X^k with the fitting of the underlying distribution with an appropriate membership function. If a membership function is treated as an agent, different forms of membership functions (agents) give different information sets for the same set of information source values. For example, in the context of the information table in Figure 1, the information associated with x_i on attribute X is given as

$$D_X(x_i) = \left[\left(I_X^1(x_i^1), \mu_X^1(x_i^1)\right), \left(I_X^2(x_i^2), \mu_X^2(x_i^2)\right), \dots, \left(I_X^z(x_i^z), \mu_X^z(x_i^z)\right)\right] \tag{8}$$

The following example would further help in illustrating the concept of information sets.

Example 3.1: Let us consider two persons A and B with salaries denoted as $x_i^{A_{high}}$ and $x_i^{B_{high}}$. Here, attribute $X =$ salary, and fuzzy sets are $k_1 = A_{high}$ and $k_2 = B_{high}$. The fuzzy sets A_{high} and B_{high} refer to the two different fuzzy sets, corresponding to A and B, for the concept high salary.

We consider the information source for x_i as $I_X^{A_{high}}\left(x_i^{A_{high}}\right) = I_X^{B_{high}}\left(x_i^{B_{high}}\right) = \$4000$. Both persons A and B have their own degree of satisfaction with their salary, reflected through their individual MFs for the fuzzy set high salary. Person A, evaluates his salary of \$4000 as relatively high and he assigns membership in the class high salary as $\mu_X^{A_{high}}\left(x_i^{A_{high}}\right) = 1$. In comparison, person B evaluates his salary not very high and assigns $\mu_X^{B_{high}}\left(x_i^{B_{high}}\right) = 0.6$.

In the original fuzzy set and other theories which extend the fuzzy sets theory like fuzzy rough sets, intuitionistic fuzzy sets or vague sets theories, the information is communicated in the form of membership grades,

$\mu_X^{A_{high}}\left(x_i^{A_{high}}\right) = 1$ or $\mu_X^{B_{high}}\left(x_i^{B_{high}}\right) = 0.6$.

In the information set concept, the information value (information source) along with an agent's judgment $\left(\mu_X^k\right)$, is considered. The information set of high salary as per A is $S_X^{A_{high}}\left(x_i^{A_{high}}\right) = I_X^{A_{high}}\left(x_i^{A_{high}}\right) * \mu_X^{A_{high}}\left(x_i^{A_{high}}\right) = \$4000(1) = \$4000$. Similarly, $S_X^{B_{high}}\left(x_i^{B_{high}}\right) = \$4000(0.6) = \$2400$.

It can easily be observed that the concept of information set agrees well with intuit. The same information source (\$4000) carries more information/value for A, who assigns a higher membership grade to it than B. In other words, an information set communicates the information as viewed by an agent.

IV. Properties of Information Sets

Here, we give a few properties of the information sets which are derived by extending the corresponding definitions for fuzzy sets and ordinary sets.

The *complement* of an information set S_X^k is defined as

$$S_X^{k^*} = \left\{I_X^k\left(x_i^k\right)\left(1 - g_X^k\left(x_i^k\right)\right)\right\}, \forall x_i^k \in k, U$$

The information of the complement is defined as

$$H_X^{k^*} = \frac{1}{|F_X^k|}\sum_i I_X^k\left(x_i^k\right)\left(1 - g_X^k\left(x_i^k\right)\right), \ \forall x_i^k \in k \qquad (9)$$

Containment: An information set $S_X^{k_1}$ is contained in $S_X^{k_2}$, that is

$$S_X^{k_1} \subseteq S_X^{k_2} \quad \text{iff} \quad g_X^{k_1}\left(x_i^{k_1}\right) \le g_X^{k_2}\left(x_i^{k_2}\right), \ \forall x_i^{k_1} \in k_1, U; \ \forall x_i^{k_2} \in k_2 \qquad (10)$$

The *union* of two information sets $S_X^{k_1}$ and $S_X^{k_2}$ with respective membership functions $g_X^{k_1}\left(x_i^{k_1}\right)$ and $g_X^{k_2}\left(x_i^{k_2}\right)$ results into an information set $S_X^{k_3} = S_X^{k_1} \cup S_X^{k_2}$. The elements of the resulting information set $S_X^{k_3}$ are related to those of $S_X^{k_1}$ and $S_X^{k_2}$ as

$$S_X^{k_3}\left(x_i^{k_3}\right) = S_X^{k_1}\left(x_i^{k_1}\right) \cup S_X^{k_2}\left(x_i^{k_2}\right) = \left\{I_X^{k_3}\left(x_i^{k_3}\right)g_X^{k_3}\left(x_i^{k_3}\right)\right\}, \forall x_i^{k_1} \in k_1; \ \forall x_i^{k_2} \in k_2; \ \forall x_i^{k_3} \in k_3 \qquad (11)$$

where, $k_3 = k_1 \cup k_2$

$g_X^{k_3}(x_i) = \max\left[g_X^{k_1}(x_i), \ g_X^{k_2}(x_i)\right]$, is the information gain for the resulting set $S_X^{k_3}$.

Theorem 4.1: The union of two $S_X^{k_1}$ and $S_X^{k_2}$ is the smallest information set containing both $S_X^{k_1}$ and $S_X^{k_2}$.

Proof:

$$\max\left[I_X^{k_1}\left(x_i^{k_1}\right)g_X^{k_1}\left(x_i^{k_1}\right), I_X^{k_2}\left(x_i^{k_2}\right)g_X^{k_2}\left(x_i^{k_2}\right)\right] \ge I_X^{k_1}\left(x_i^{k_1}\right)g_X^{k_1}\left(x_i^{k_1}\right)$$

$$\max\left[I_X^{k_1}\left(x_i^{k_1}\right)g_X^{k_1}\left(x_i^{k_1}\right), I_X^{k_2}\left(x_i^{k_2}\right)g_X^{k_2}\left(x_i^{k_2}\right)\right] \ge I_X^{k_2}\left(x_i^{k_2}\right)g_X^{k_2}\left(x_i^{k_2}\right)$$

Furthermore, if S_X^{k} is any information set containing both $S_X^{k_1}$ and $S_X^{k_2}$, then from (11),

$$I_X^{k}\left(x_i^{k}\right)g_X^{k}\left(x_i^{k}\right) \ge I_X^{k_1}\left(x_i^{k_1}\right)g_X^{k_1}\left(x_i^{k_1}\right), \forall x_i \in k_1, U$$
$$I_X^{k}\left(x_i^{k}\right)g_X^{k}\left(x_i^{k}\right) \ge I_X^{k_2}\left(x_i^{k_2}\right)g_X^{k_2}\left(x_i^{k_2}\right), \forall x_i \in k_2, U$$

Hence,

$$I_X^{k}\left(x_i^{k}\right)g_X^{k}\left(x_i^{k}\right) \ge \max\left[I_X^{k_1}\left(x_i^{k_1}\right)g_X^{k_1}\left(x_i^{k_1}\right), I_X^{k_2}\left(x_i^{k_2}\right)g_X^{k_2}\left(x_i^{k_2}\right)\right] = I_X^{k_3}\left(x_i^{k_3}\right)g_X^{k_3}\left(x_i^{k_3}\right)$$

In light of above, $S_X^{k_3} \le S_X^{k}$ follows.

The *intersection* of two information sets $S_X^{k_1}$ and $S_X^{k_2}$ with respective membership functions $g_X^{k_1}\left(x_i^{k_1}\right)$ and $g_X^{k_2}\left(x_i^{k_2}\right)$ results into an information set $S_X^{k_3}$, denoted as $S_X^{k_3} = S_X^{k_1} \cap S_X^{k_2}$. The elements of the resulting information set $S_X^{k_3}$ are related to those of $S_X^{k_1}$ and $S_X^{k_2}$ as

$$S_X^{k_3}\left(x_i^{k_3}\right) = S_X^{k_1}\left(x_i^{k_1}\right) \cap S_X^{k_2}\left(x_i^{k_2}\right) = \left\{I_X^{k_3}\left(x_i^{k_3}\right)g_X^{k_3}\left(x_i^{k_3}\right)\right\}, \forall x_i^{k_1} \in k_1; \ \forall x_i^{k_2} \in k_2; \ \forall x_i^{k_3} \in k_3 \qquad (12)$$

where, $g_X^{k_3}\left(x_i^{k_3}\right) = \min\left[g_X^{k_1}\left(x_i^{k_1}\right), \ g_X^{k_2}\left(x_i^{k_2}\right)\right]$ is the information gain for the resulting set $S_X^{k_3}$.

Theorem 4.2: With the operations of union and intersection defined in (11) and (12), the idempotency, commutative, associative and distributive properties, and De Morgan's laws hold true.

Proof: The proofs for idempotency and commutative properties follow trivially as follows.

Idempotency: $S_X^{k_1} \cup S_X^{k_1} = S_X^{k_1}$ (13)

From (11),

$$S_X^{k_1}\left(x_i^{k_1}\right) \cup S_X^{k_1}\left(x_i^{k_1}\right) = \max\left[I_X^{k_1}\left(x_i^{k_1}\right)g_X^{k_1}\left(x_i^{k_1}\right), I_X^{k_1}\left(x_i^{k_1}\right)g_X^{k_1}\left(x_i^{k_1}\right)\right], \forall x_i^{k_1} \in k_1, U = S_X^{k_1}$$

Analogously, $S_X^{k_1} \cap S_X^{k_1} = S_X^{k_1}$ (14)

Commutativity: $S_X^{k_1} \cup S_X^{k_2} = S_X^{k_2} \cup S_X^{k_1}$ (15)

Since,

$$\max\left[I_X^{k_1}(x_i^{k_1})g_X^{k_1}(x_i^{k_1}), I_X^{k_2}(x_i^{k_2})g_X^{k_2}(x_i^{k_2})\right] = \max\left[I_X^{k_2}(x_i^{k_2})g_X^{k_2}(x_i^{k_2}), I_X^{k_1}(x_i^{k_1})g_X^{k_1}(x_i^{k_1})\right]$$

the commutative property holds good for information sets.

Similarly, the commutative property stands true for intersection operation as well.

Associativity: $S_X^{k_1} \cup (S_X^{k_2} \cup S_X^{k_3}) = (S_X^{k_1} \cup S_X^{k_2}) \cup S_X^{k_3}$ (16)

LHS: From (11), $S_X^{k_1}(x_i^{k_1}) \cup \left(S_X^{k_2}(x_i^{k_2}) \cup S_X^{k_3}(x_i^{k_3})\right)$

$$= \max\begin{bmatrix} I_X^{k_1}(x_i^{k_1})g_X^{k_1}(x_i^{k_1}), \\ \max\left[I_X^{k_2}(x_i^{k_2})g_X^{k_2}(x_i^{k_2}), I_X^{k_3}(x_i^{k_3})g_X^{k_3}(x_i^{k_3})\right] \end{bmatrix}$$

$$= \max\left[I_X^{k_1}(x_i^{k_1})g_X^{k_1}(x_i^{k_1}), I_X^{k_2}(x_i^{k_2})g_X^{k_2}(x_i^{k_2}), I_X^{k_3}(x_i^{k_3})g_X^{k_3}(x_i^{k_3})\right]$$

RHS: similarly,

$$\left(S_X^{k_1}(x_i^{k_1}) \cup S_X^{k_2}(x_i^{k_2})\right) \cup S_X^{k_3}(x_i^{k_3})$$

$$= \max\left[I_X^{k_1}(x_i^{k_1})g_X^{k_1}(x_i^{k_1}), \quad I_X^{k_2}(x_i^{k_2})g_X^{k_2}(x_i^{k_2}), I_X^{k_3}(x_i^{k_3})g_X^{k_3}(x_i^{k_3})\right]$$

Hence,

$$S_X^{k_1} \cup (S_X^{k_2} \cup S_X^{k_3}) = (S_X^{k_1} \cup S_X^{k_2}) \cup S_X^{k_3}$$

The proof for associative property for intersection operation follows analogously.

Distributivity:

$$S_X^{k_1} \cap (S_X^{k_2} \cup S_X^{k_3}) = (S_X^{k_1} \cap S_X^{k_2}) \cup (S_X^{k_1} \cap S_X^{k_3}) \quad (17)$$

Consider the LHS: $S_X^{k_1} \cap (S_X^{k_2} \cup S_X^{k_3})$.

Let,

$$x \in S_X^{k_1} \cap (S_X^{k_2} \cup S_X^{k_3}) \Rightarrow x \in S_X^{k_1} \wedge x \in (S_X^{k_2} \cup S_X^{k_3})$$

$$\Rightarrow x \in S_X^{k_1} \wedge \{x \in S_X^{k_2} \vee x \in S_X^{k_3}\}$$

$$\Rightarrow \{x \in S_X^{k_1} \wedge x \in S_X^{k_2}\} \vee \{x \in S_X^{k_1} \wedge x \in S_X^{k_3}\}$$

$$\Rightarrow \{x \in (S_X^{k_1} \cap S_X^{k_2})\} \vee \{x \in (S_X^{k_1} \cap S_X^{k_3})\}$$

The proof follows trivially as in the ordinary sets.

Analogously,

it is easy to show that

$$S_X^{k_1} \cup (S_X^{k_2} \cap S_X^{k_3}) = (S_X^{k_1} \cup S_X^{k_2}) \cap (S_X^{k_1} \cup S_X^{k_3}) \quad (18)$$

De Morgan's laws

$$\left(S_X^{k_1}(x_i^{k_1}) \cup S_X^{k_2}(x_i^{k_2})\right)^* = S_X^{k_1}(x_i^{k_1})^* \cap S_X^{k_2*}(x_i^{k_2}) \quad (19)$$

Consider LHS: $\left(S_X^{k_1}(x_i^{k_1}) \cup S_X^{k_2}(x_i^{k_2})\right) =$

$$\max\left(I_X^{k_1}(x_i^{k_1})g_X^{k_1}(x_i^{k_1}), \quad I_X^{k_2}(x_i^{k_2})g_X^{k_2}(x_i^{k_2})\right)$$

Suppose, $I_X^{k_1}(x_i^{k_1})g_X^{k_1}(x_i^{k_1}) \geq I_X^{k_2}(x_i^{k_2})g_X^{k_2}(x_i^{k_2})$

Then,

$$\left(S_X^{k_1}(x_i^{k_1}) \cup S_X^{k_2}(x_i^{k_2})\right)^* = I_X^{k_1}(x_i^{k_1})[1 - g_X^{k_1}(x_i^{k_1})]$$

Consider RHS:

$$S_X^{k_1*}(x_i^{k_1}) \cap S_X^{k_2*}(x_i^{k_2}) = min\left[I_X^{k_1}(x_i^{k_1})\left(1 - g_X^{k_1}(x_i^{k_1})\right), I_X^{k_2}(x_i^{k_2})\left(1 - g_X^{k_2}(x_i^{k_2})\right)\right]$$

Since,

$$I_X^{k_1}(x_i^{k_1})g_X^{k_1}(x_i^{k_1}) \geq I_X^{k_2}(x_i^{k_2})g_X^{k_2}(x_i^{k_2})$$

$$\Rightarrow g_X^{k_1}(x_i^{k_1}) \geq g_X^{k_2}(x_i^{k_2}) \quad (20)$$

In light of (12), RHS of (20) yields

$$S_X^{k_1*}(x_i^{k_1}) \cap S_X^{k_2*}(x_i^{k_2}) = min\left[I_X^{k_1}(x_i^{k_1})\left(1 - g_X^{k_1}(x_i^{k_1})\right), I_X^{k_2}(x_i^{k_2})\left(1 - g_X^{k_2}(x_i^{k_2})\right)\right]$$

$$= I_X^{k_1}(x_i^{k_1})\left(1 - g_X^{k_1}(x_i^{k_1})\right) = \text{LHS}.$$

Similarly, it can be shown that $(S_X^{k_1} \cap S_X^{k_2})^* = S_X^{k_1*} \cup S_X^{k_2*}$.

V. Information Measures for Information Sets

In this section, a few measures for the information sets have been developed for the information sets for a given information table $T = (U, At, \{I_X | X \in At\})$.

A. The Information Measures

Definition 5.1: The net information quantity for a data element x_i over all the fuzzy sets is given as

$$N_X(x_i) = -\sum_k I_X(x_i)\log S_X^k(x_i) \quad (21)$$

It gives the net information quantity for a data element x_i summed over all the fuzzy sets. All such values of $N_X(x_i)$ constitute a vector N_X.

Definition 5.2: The net information represented by information table T is given as

$$NH_X = -\sum_{i=1}^{n} N_X(x_i) \ \log N_X(x_i) \ , \forall x_i \in U \quad (22)$$

Definition 5.3: The net divergence between various elements of an information set is given as

$$DH_X = -\sum_i D_X(x_i)\log D_X(x_i)\,, \forall x_i \in U \quad (23)$$

The more the value of DH_X, more is the net dispersion between the elements $\{x_i\}$ of information source IS_X.

Definition 5.4: The relative information $D_X^{k_1/k_2}$ for the fuzzy sets/concepts k_1 and k_2, defined for all $x_i \epsilon U$ is given as

$$D_X^{k_1/k_2} = -\sum_{i=1}^{n} \mu_{k_1}(x_i)log\frac{S_X^{k_1}(x_i)}{S_X^{k_2}(x_i)}, \forall x_i \in U \quad (24)$$

where, $S_X^{k_1}(x_i) = I_X(x_i)\mu_{k_1}(x_i)$; $S_X^{k_2}(x_i) = I_X(x_i)\mu_{k_2}(x_i)$

Definition 5.5: The joint information for an attribute X over fuzzy sets k_1 and k_2 can be computed as

$$D_X^{k_1k_2} = -\sum_{i=1}^{n} \mu_{k_1}(x_i)\mu_{k_2}(x_i)\, log\, S_X^{k_1}(x_i)S_X^{k_2}(x_i) \tag{25}$$

Similarly, it can be shown that the other measures also hold true for the non-linear information sets, as well.

B. Interpretation of information Measures with fuzzy sets

We now take a few examples of fuzzy sets and derive the corresponding information sets and their measures.

Example 5.1

We consider a triangular fuzzy set as given in Figure 3 for a population database giving the daily wages X in \$. The objects in the population database are denoted as $\{x_i^1\}_{i=1}^n$.

Let us consider the information source, $I_X^1(x_i^1) = \{100, 200, 300, 400, 500\}$, as shown in Figure 3.

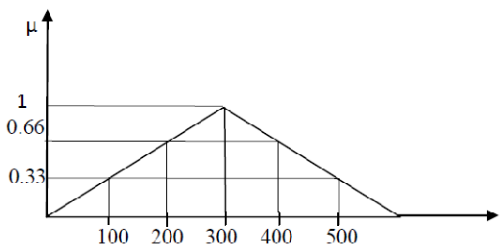

Fig. 3: A Triangular Fuzzy Set 1

The fuzzy set corresponding to Fig. 4 is $\{0.33,0.66,1,0.66,0.33\}$. The corresponding information set is given as $S_X^1 = \{33,132,300,264,165\}$. The information for S_X^1 is given as

$$H_X^1 = \frac{1}{5}(33 + 132 + 300 + 264 + 165) = 178.80$$

It can be seen that the fuzzy set is only reporting the membership grades, while the information set represents information, as viewed by the agent. Now, we study the effect of repeated occurrences of values in the database. Let $I_X^1(x_i^1) =$ $\{100, 200, 200,200, 300,300, 400, 400,400, 400, 400, 500\}$

The information set is given as

$S_X^1 = \{33,132,300,300,300,264,264,264,264,264,165\}$

The information for S_X^1 is

$$H_X^1 = \frac{1}{12}\big((1)(100)(0.33) + (3)(200)(0.66) + (2)(300)(1) + (5)(400)(0.66) + (1)(500)(0.33)\big) = 195.75$$

Example 5.2

Here, a combination of fuzzy sets is considered, as shown in Figure 4. We take a population database, of monthly wages in \$. Here, C_1 = Low, C_2 = Medium, C_3 = High. The information sources are $I_X^{C_1}(x_i^{C_1}) = \{1000, 3000, 5000\}$, $I_X^{C_2}(x_i^{C_2}) = \{1000, 3000, 5000, 7000, 10000\}$, $I_X^{C_2}(x_i^{C_2}) =$ $\{5000, 7000, 10000\}$

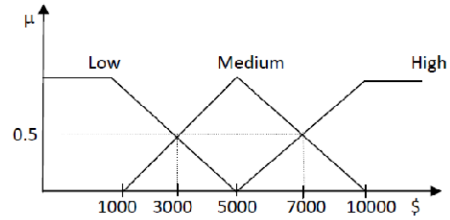

Fig. 4: An Example for Combination of Fuzzy Sets

The information sets corresponding to fuzzy sets **Low**, **Medium** and **High** are given as follows

$S_X^{C_1} = \{1000, 1500, 0\}$ $H_X^{C_1} =$ 833.33

$S_X^{C_2} = \{1000,1500,5000,3500\}$, $H_X^{C_2} = 2750$

$S_X^{C_3} = \{0, 3500, 10000\}$ $H_X^{C_3} = 4500$

Example 5.3

In this example, we compute the information measures for information sets. We consider a combination of two fuzzy sets, Low and High, as shown in Figure 5, and a population database, of monthly wages in \$.

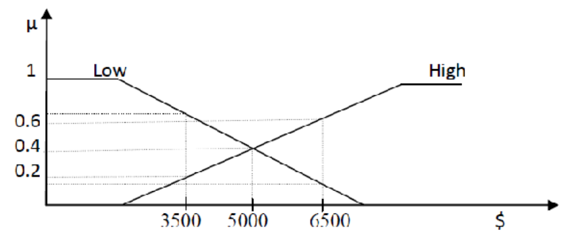

Fig. 5: An Example for Combination of Fuzzy Sets

The fuzzy sets Low and High in Figure 5 give the qualitative representation of attribute X, monthly **wages**.

Here, $C_1 = Low$, $C_2 = High$, $I_X^{C_1}(x_i^{C_1}) = I_X^{C_2}(x_i^{C_2}) =$ $\{3500, 5000, 6500\}$.

The information sets corresponding to C_1 and C_2 are given as follows

$S_X^{C_1} = \{2170, 2000,1170\}$; $S_X^{C_2} = \{700, 2000, 3900\}$

The information values corresponding to $S_X^{C_1}$ and $S_X^{C_2}$ are given as follows

$H_X^{C_1} = 1780$; $H_X^{C_2} = 2200$

We, now compute the information measures for the information sets with $\{x_i\} = \{3500, 5000, 6500\}$, as follows.

The relative information $D_X^{Low/High}$ for the fuzzy sets **Low** and **High** is given as

$$D_X^{Low/High} = -\big(0.62log(3.1) + 0.4log(1) + 0.18log(0.3)\big) = -0.4848$$

$$D_X^{High/Low} = 0.20log(0.32) + 0.4log(1) + 0.60log(3.33) = -0.4939$$

The joint information for attribute X over fuzzy sets **Low** and **High,** $D_X^{High.Low}$ is given as

$$D_X^{High.Low} = -\big((0.62)(0.20)log3500(0.62)3500(0.20) + (0.4)(0.4)log5000(0.4)5000(0.4) + (0.18)(0.60)\log 6500(0.18)6500(0.60)\big) = -5.8583$$

The net information quantity for various data elements over all the fuzzy sets is given as

$$N_X(3500) = -(3500\log(2170) + 3500\log(700)) = -4.9817e+04$$
$$N_X(5000) = -(5000\log(2000) + 5000\log(2000)) = -7.6009e+04$$
$$N_X(6500) = -(6500\log(3900) + 6500\log(1170)) = -9.9668e+04$$

The net information represented by the information table T is given as

$$NH_X = -225494$$

The information source values are normalized as $\left(I_X^{C_j}\left(x_i^{C_j}\right)\right)_N = \frac{I_X^{C_j}\left(x_i^{C_j}\right) - I_X^{C_j}(x_{min})}{I_X^{C_j}(x_{max}) - I_X^{C_j}(x_{min})}$. The normalized values are obtained as $\{0, 0.5, 1\}$.

The relative information $D_X^{C_1/C_2}$ for the fuzzy sets **Low** and **High** is given as

$$D_X^{C_1/C_2}(0) = -\left(0.62log\left(\frac{0*0.62}{1*0.18}\right) + 0.62log\left(\frac{0*0.62}{0.5*0.4}\right)\right) = \varepsilon$$
$$D_X^{C_1/C_2}(0.5) = -\left(0.4log\left(\frac{0.5*0.4}{1*0.18}\right) + 0.4log\left(\frac{0.5*0.4}{0*0.62}\right)\right) = -0.0422 + \varepsilon$$
$$D_X^{C_1/C_2}(1) = -\left(0.18log\left(\frac{1*0.18}{0.5*0.4}\right) + 0.18log\left(\frac{1*0.18}{0*0.62}\right)\right) = 0.0190 + \varepsilon$$
$$D_X^{C_2/C_1}(0) = -\left(0.20log\left(\frac{0*0.20}{1*0.60}\right) + 0.20log\left(\frac{0*0.20}{0.5*0.4}\right)\right) = \varepsilon$$
$$D_X^{C_2/C_1}(0.5) = -\left(0.4log\left(\frac{0.5*0.4}{1*0.60}\right) + 0.4log\left(\frac{0.5*0.4}{0*0.20}\right)\right) = 0.4394 + \varepsilon$$
$$D_X^{C_2/C_1}(1) = -\left(0.60log\left(\frac{1*0.60}{0.5*0.4}\right) + 0.60log\left(\frac{1*0.60}{0*0.20}\right)\right) = -0.6592 + \varepsilon$$

VI. Conclusions

The properties of the information sets are investigated in detail. It is found that most of the standard properties hold good with the information sets like the case with the classical sets. The information measures are also developed for the information sets. The usefulness of the information measures in computing the information associated with the information sets is also shown through a set of three illustrative examples.

References

[1] N. D. Singpurwalla, J. M. Booker, Membership Functions and Probability Measures of Fuzzy Sets. J. Am. Stat. Assoc., 99 (467), pp. 867–877, 2004

[2] M. Agarwal, M. Hanmandlu, K. K. Biswas, Information Sets, IEEE Transactions on Fuzzy Systems. Communicated.

[3] Z. Pawlak, Rough Sets, International Journal of Computer and Information Sciences 11 (1982) 341-356.

[4] L.A. Zadeh, Forward, in: E. Orlowska (Ed.), Incomplete Information: Rough Set Analysis, Physica-Verlag, Heidelberg, 1998, pp. v-vi.

[5] Yao, Y.Y., A comparative study of fuzzy sets and rough sets, Information Sciences, Vol. 109, No. 1-4, pp. 227-242, 1998.

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

[7] P. Nowak, M. Romaniuk, Application of Levy processes and Esscher transformed martingale measures for option pricing in fuzzy framework, Journal of Computational and Applied Mathematics, Vol. 263, Jun. 2014, Pp. 129-151.

[8] L. Zhou, Z. Tao, H. Chen, J. Liu, Continuous interval-valued intuitionistic fuzzy aggregation operators and their applications to group decision making, Applied Mathematical Modelling, Vol. 38, 7–8, Apr 2014, Pp. 2190-2205

[9] S. Kar, S. Das, P. K. Ghosh, Applications of neuro fuzzy systems: A brief review and future outline, Applied Soft Computing, Vol. 15, Feb. 2014, Pp. 243-259

[10] C. Bai, D. Dhavale, J. Sarkis, Integrating Fuzzy C-Means and TOPSIS for performance evaluation: An application and comparative analysis, Expert Systems with Applications, Vol. 41, 9, Jul. 2014, pp 4186-4196,

[11] A. Beucher, S. Fröjdö, P. Österholm, A. Martinkauppi, P. Edén, Fuzzy logic for acid sulfate soil mapping: Application to the southern part of the Finnish coastal areas, Geoderma, Vol. 226–227, Aug. 2014, pp. 21-30.

[12] J. Lawry, Probability, fuzziness and borderline cases, International Journal of Approximate Reasoning, Available online 28 Jan. 2014.

[13] J. Dombi, Basic concepts for a theory of evaluation: The aggregative operator, European Journal of Operational Research, Vol. 10, 3, July 1982, Pp. 282-293.

[14] J. Lawry, Y. Tang, Granular Knowledge Representation and Inference Using Labels and Label Expressions, Fuzzy Systems, IEEE Transactions on , vol.18, no.3, pp.500,514, June 2010