

# Probabilistic Fuzzy Reasoning in a Coherent Setting

Giulianella Coletti<sup>1</sup> Davide Petturiti<sup>2</sup> Barbara Vantaggi<sup>2</sup>

<sup>1</sup>Dip. Matematica e Informatica, University of Perugia, Italy

<sup>2</sup>Dip. S.B.A.I., “La Sapienza” University of Rome, Italy

## Abstract

We use the interpretation of fuzzy sets in terms of coherent conditional probabilities for handling probabilistic fuzzy IF-THEN rules. We show by some examples how this interpretation can help when fuzzy and statistical information need to be combined and the available probabilistic information on the fuzzy events is possibly imprecise or incomplete.

**Keywords:** Probabilistic Fuzzy Reasoning, Probabilistic Fuzzy IF-THEN Rule Base, Coherent Conditional Probability

## 1. Introduction

In the last decades the fuzzy set theory has seen a growing interest in the literature both from a practical and theoretical point of view. One of the main objectives is to jointly handle probability and vagueness [3, 4, 6, 8, 9, 11, 14, 15, 16, 21, 22, 23, 31, 32, 39, 40, 41, 46, 43, 44, 47, 48, 49]. For this task it is suitable to have a general framework in which it is possible to handle uncertainty encoding the different aspect by maintaining consistency with the model of reference: this is crucial to ensure a sound inference.

In this paper we refer to the interpretation of a fuzzy subset  $E_\varphi^*$  as a pair  $(E_\varphi, \mu_\varphi)$  with  $\mu_\varphi(x) = P(E_\varphi|X = x)$ , where  $X$  is a variable ranging in the set  $\mathcal{C}_X$ ,  $\varphi$  is any property related to  $X$ ,  $E_\varphi$  is the Boolean event “You claim that  $X$  is  $\varphi$ ” and  $P$  is a coherent conditional probability. We are interested in evaluating the (conditional) probability of “fuzzy events”.

In the literature, the notion of probability has been extended to non-Boolean events represented as elements of a MV-algebra [35]. In this environment, conditioning and coherence have been also investigated [36, 34] obtaining to some extent a many-valued generalization of existing concepts.

Our aim is not to endow a set of fuzzy events with a predetermined algebraic structure induced by a predetermined t-norm (and t-conorm) but rather to show how our interpretation is useful when we have to deal with fuzzy and probabilistic information. In particular the paper focuses on probabilistic fuzzy reasoning based on collections of probabilistic fuzzy

IF-THEN rules, showing their applicability to general problems such as fuzzy classification and fuzzy control.

In practical situations, the available information is often partial or even only available on sets of events which are not those of interest and come out possibly from different sources. Then the general framework of coherent conditional probability provides accurate tools for testing the coherence (i.e., the consistency with the model of reference) of the available assignments and for ruling their extension to other events of interest (that is for making inference). Usually the extension is not unique and so in order to preserve as much as possible the information at hand we must refer to the whole set of coherent extensions and so to coherent lower and upper conditional probabilities.

Concerning coherent lower and upper conditional probabilities often their computation is not easy and there are no explicit analytic formulas. On the other hand, in some cases these uncertainty measures have features which allow to obtain the desired goals by a direct computation, for instance by using the Choquet integral (see Example 2). Other difficulties arise when events with zero lower probability are involved, in fact unconditional measures are not able to bound the extensions that can range in the entire interval  $[0, 1]$ , and so a conditional assessment needs to be considered [13] or otherwise some specific restriction has to be adopted [33].

By using coherent conditional probability as unifying model for every kind of uncertainty (produced by different sources, but essentially due to incomplete information) makes possible to treat probabilistic fuzzy IF-THEN rules possibly involving fuzzy subsets both in the antecedent and in the consequent. As a side effect, it is possible to benefit of some existing results related to other reasoning models connected with coherent conditional probabilities such as *System P* or *default reasoning* for Boolean events (see for instance [27, 28, 29, 12, 11, 15]).

This is a position paper focusing on the analysis of some paradigmatic examples, in a way to show the potentiality of the coherent conditional probability framework for encoding and managing fuzzy information. The theoretical results of reference are just recalled in the Appendix A since they are essentially developed in previous papers.

## 2. Probabilistic Fuzzy IF-THEN rules

A relevant problem in literature is the managing of fuzzy IF-THEN rules based systems: they are essentially composed by a collection of rules of the form

“IF  $A$  THEN  $B$ , with a given probability”,

where either premise  $A$  and consequence  $B$  of the rule can be fuzzy subsets [4, 31, 32, 39, 41, 46, 49]. Here we study this problem by adopting the interpretation of fuzzy sets in terms of coherent conditional probabilities introduced in [7, 8, 9] (and developed in [40, 6, 14, 15]) whose main concepts are recalled in Appendix A. Under this interpretation the membership function  $\mu_\varphi(\cdot)$  of a fuzzy set related to the property  $\varphi$  is represented by a conditional probability  $P(E_\varphi|\cdot)$  regarded as function of the conditioning event.

A noticeable example in the above systems is the task of determining the class  $Y \in \{c_1, \dots, c_s\}$  to which a data point  $x = (x_1, \dots, x_m)$  related to the vector  $X = (X_1, \dots, X_m)$  belongs. For this aim a probabilistic fuzzy classifier could be performed, by giving some rules such as:

IF  $X$  IS  $\varphi$  THEN  $Y = c_k$ , with probability  $p_{k|\varphi}$ ,

where “ $X$  IS  $\varphi$ ” stands for “You claim that  $X$  IS  $\varphi$ ” (see Appendix A). By using these probabilistic rules between crisp or fuzzy events, a related task is to determine the “most probable” point  $x$  under the hypothesis that  $Y$  has class  $c_k$ .

More generally, given a finite collection of fuzzy subsets  $\mathcal{E}^* = \{E_{\varphi_j}^* = (E_{\varphi_j}, \mu_{\varphi_j})\}_{j=1, \dots, m}$ , related to the components of the vector  $X$ , with  $\mu_{\varphi_j}(x_j) = P(E_{\varphi_j}|x_j)$  for any  $x_j \in \mathcal{C}_{X_j}$ , let us consider the fuzzy set  $E_\varphi^* = (E_\varphi, \mu_\varphi)$  obtained as intersection of some  $E_{\varphi_j}^*$ ’s (see Appendix A for details on fuzzy union and intersection). The problem consists in determining

$$P(A|E_\varphi),$$

where  $A$  can be either an event of the Boolean algebra  $\langle \mathcal{C}_Y \rangle$  generated by  $\mathcal{C}_Y$  or a fuzzy event  $E_\psi$  related to a fuzzy subset  $E_\psi^*$  of  $\mathcal{C}_Y$ .

The main problem related to the performance of a *probabilistic fuzzy classifier* (i.e., when  $A = c_r$ ) consists into the elicitation of the probabilities  $p_{k|\varphi}$ :

$$p_{k|\varphi} = P(c_k|E_\varphi) = \frac{P(E_\varphi|c_k)P(c_k)}{\sum_{r=1}^s P(E_\varphi|c_r)P(c_r)}.$$

However, when the probability distribution on the  $c_k$ ’s is not completely available or the membership function is defined on a different partition (even if related to that of the  $c_k$ ’s), the unicity of  $p_{k|\varphi}$  is not guaranteed, so the lower and upper envelope of a class  $\{p_{k|\varphi}\}$  of such probabilities must be computed.

The simplest case is when a joint probability on the vector  $X$  is available: the following trivial example shows a probabilistic fuzzy IF-THEN rule base

for the humidity forecasting of a garden center’s refrigerating room, given the temperature.

**Example 1.** Denote with  $X$  and  $Y$ , respectively, the temperature and the relative humidity of a garden center’s refrigerating room. The temperature of the system is guaranteed to range from a nominal minimum of  $0^\circ\text{C}$  to a nominal maximum of  $20^\circ\text{C}$ , thus we can put  $\mathcal{C}_X = [0, 20]$  and  $\mathcal{C}_Y = [0, 1]$ . After an empirical analysis, the joint probability distribution of  $(X, Y)$  has been approximated with the joint density function

$$f(x, y) = \begin{cases} \frac{3}{8030}(x^2 + y), & (x, y) \in [0, 20] \times [0, 1], \\ 0, & \text{otherwise.} \end{cases}$$

For the variable  $X$ , consider the linguistic properties  $\varphi_1 = \text{“cold”}$ ,  $\varphi_2 = \text{“warm”}$  and  $\varphi_3 = \text{“hot”}$ , and the relevant fuzzy subsets  $E_{\varphi_i}^*$ ,  $i = 1, 2, 3$ , of  $\mathcal{C}_X$ , having the following membership functions

$$\begin{aligned} \mu_{\varphi_1}(x) &= \begin{cases} 1, & \text{if } x < 2.5, \\ -\frac{x}{5} + \frac{3}{2}, & \text{if } x \in [2.5, 7.5], \\ 0, & \text{if } x > 7.5, \end{cases} \\ \mu_{\varphi_2}(x) &= \begin{cases} 0, & \text{if } x < 2.5, \\ \frac{x}{5} - \frac{1}{2}, & \text{if } x \in [2.5, 7.5], \\ 1, & \text{if } x \in [7.5, 12.5], \\ -\frac{x}{5} + \frac{7}{2}, & \text{if } x \in (12.5, 17.5], \\ 0, & \text{if } x > 17.5, \end{cases} \\ \mu_{\varphi_3}(x) &= \begin{cases} 0, & \text{if } x < 12.5, \\ \frac{x}{5} - \frac{5}{2}, & \text{if } x \in [12.5, 17.5], \\ 1, & \text{if } x > 17.5. \end{cases} \end{aligned}$$

In analogy, for the variable  $Y$ , consider the linguistic properties  $\psi_1 = \text{“dry”}$  and  $\psi_2 = \text{“wet”}$ , to which we relate the fuzzy subsets of  $\mathcal{C}_Y$ ,  $E_{\psi_j}^*$ ,  $j = 1, 2$ , having the following membership functions

$$\begin{aligned} \mu_{\psi_1}(y) &= \begin{cases} 1, & \text{if } y < 0.3, \\ -10y + 4, & \text{if } y \in [0.3, 0.4], \\ 0, & \text{if } y > 0.4, \end{cases} \\ \mu_{\psi_2}(y) &= \begin{cases} 0, & \text{if } y < 0.3, \\ 10y - 3, & \text{if } y \in [0.3, 0.4], \\ 1, & \text{if } y > 0.4. \end{cases} \end{aligned}$$

Choosing the algebraic product  $T_P$  as  $t$ -norm (that is a coherent choice, see Appendix A), it is possible to compute the coherent  $T_P$ -extension of the above globally coherent assessment. In particular, we have  $P_{T_P}(E_{\varphi_1}) = 0.0204$ ,  $P_{T_P}(E_{\varphi_2}) = 0.4144$  and  $P_{T_P}(E_{\varphi_3}) = 0.5652$ .

Thus, a probabilistic fuzzy IF-THEN rule base deriving from the problem is the following one:

- IF  $X$  IS cold THEN
  - $Y$  IS dry,  $P_{T_P}(E_{\psi_1}|E_{\varphi_1}) = 0.3396$ ;
  - $Y$  IS wet,  $P_{T_P}(E_{\psi_2}|E_{\varphi_1}) = 0.6604$ ;
- IF  $X$  IS warm THEN
  - $Y$  IS dry,  $P_{T_P}(E_{\psi_1}|E_{\varphi_2}) = 0.3490$ ;
  - $Y$  IS wet,  $P_{T_P}(E_{\psi_2}|E_{\varphi_2}) = 0.6510$ ;

• IF  $X$  IS hot THEN

- $Y$  IS dry,  $P_{T_P}(E_{\psi_1}|E_{\varphi_3}) = 0.3496$ ;
- $Y$  IS wet,  $P_{T_P}(E_{\psi_2}|E_{\varphi_3}) = 0.6504$ .

Notice that probabilities  $P_{T_P}(E_{\varphi_i})$ ,  $i = 1, 2, 3$ , as well as probabilities  $P_{T_P}(E_{\psi_j}|E_{\varphi_i})$ ,  $j = 1, 2$ , for every given  $E_{\varphi_i}$ , in this particular case, sum up to 1. This holds since the product  $t$ -norm is chosen and  $\{E_{\varphi_i}^*\}_{i=1,2,3}$  and  $\{E_{\psi_j}^*\}_{j=1,2}$  are fuzzy partitions in the sense of Ruspini (i.e., for every  $x \in \mathcal{C}_X$ ,  $\sum_{i=1}^3 \mu_{\varphi_i}(x) = 1$  and for every  $y \in \mathcal{C}_Y$ ,  $\sum_{j=1}^2 \mu_{\psi_j}(y) = 1$ ). In general the sum of  $P_{T_P}(E_{\varphi_i})$ ,  $i = 1, 2, 3$  (as  $P_{T_P}(E_{\psi_j}|E_{\varphi_i})$ ,  $j = 1, 2$ ) need not be 1 for arbitrary fuzzy subsets or a different  $t$ -norm.

Until now we considered the case where the probability distribution on the variable  $X$ , to which the fuzzy subset refers, is available. However in many situations this condition does not hold (see, e.g., [6]): consider, for instance, to evaluate the uncertainty on the elements of a database (i.e., a partition) having statistical information elicited on a set of values related to some feature and a membership related to a property on a different (but related) attribute. First of all the global coherence of the whole assessment needs to be checked, and if it holds, then by Theorem 1 in Appendix A we can extend the whole assessment to other elements.

Nevertheless, as the extension is generally non-unique, we get a set of conditional probabilities and so the coherent lower (and upper) conditional probability. A particular but interesting case is when the lower probability is 2-monotone, i.e., for any  $A, B$  in the Boolean algebra  $\langle \mathcal{C}_X \rangle$ ,

$$\underline{P}(A \vee B) \geq \underline{P}(A) + \underline{P}(B) - \underline{P}(A \wedge B).$$

In this case, the infimum of the expected values of a bounded function on  $\mathcal{C}_X$  computed w.r.t. a class of probabilities is obtained by the Choquet integral w.r.t. the lower probability  $\underline{P}$  (see e.g. [20]). Then, the value  $\underline{P}(E_{\varphi})$  can be directly computed as the Choquet integral of the membership  $\mu_{\varphi}$  w.r.t.  $\underline{P}$ . More precisely, being  $\mu_{\varphi}(\cdot) = P(E_{\varphi}|\cdot)$  a function with values in  $[0, 1]$

$$\oint \mu_{\varphi} d\underline{P} = \int_0^1 \underline{P}(\{s : \mu_{\varphi}(s) \geq x\}) dx.$$

Notice that the 2-monotonicity property is verified, for instance, when the lower probability on a partition (or Boolean algebra) is obtained as lower envelope of the coherent extensions of a probability defined on a different partition (or Boolean algebra): in this case in fact, independently of the logical constraints among the events of the two partitions we obtain a completely monotone (and so 2-monotone) lower probability (i.e., a belief function, see, e.g., [25, 10]). This situation applies when the fuzzy set is defined on a partition different from that one where the probability distribution is defined. Therefore,

in these cases

$$\{\underline{P}_T(E_{\varphi_i}), \underline{P}_T(E_{\varphi_i} \wedge E_{\varphi_j})\}$$

can be computed by means of the Choquet integral.

A different situation arises when we extend on a Boolean superalgebra several probabilities assessed on different Boolean algebras. This happens, for instance, when we have marginal probabilities and we are interested in obtaining the joint probability (on the conjunctions). In this case in fact we could have a non-2-monotone extension (see [14] for an example).

An interesting case in statistics is related to imputation and statistical matching [38, 17], where the information come from different sources. These two problems have been studied in terms of coherence in [1, 42], respectively. To give an example of statistical matching, consider three categorical variables  $A$  = “age”,  $EL$  = “educational level” and  $PS$  = “professional status”, which are not jointly observed. In file 1, the variables  $A$  and  $PS$  are observed, while file 2 the variables  $A$  and  $EL$  are considered.

Logical constraints between the variables  $A$  and  $EL$  (as well as, between  $A$  and  $PS$ ) are present: for example, in Italy a 17 years old person cannot have a University degree. By considering the maximum likelihood estimations as evaluation of the relevant conditional probabilities, due to logical constraints, the global assessment could be not coherent (see [42]), then the assessment has to be corrected with respect to a pseudo-distance [2] and so the corrected assessment could not be 2-monotone.

Now, by considering the fuzzy subsets related to  $EL$  (“high”, “medium” and “low”) and to  $PS$  (“good”, “middle”, “bad”) we could look for the evaluation of  $EL$  is “high” and  $PS$  is “good”. When the lower (conditional) probability on  $EL$  and  $PS$  (given  $A$ ) is 2-monotone this evaluation can be computed by means of Choquet integral otherwise we need to find, for any coherent extension, the probability of the relevant fuzzy event and then take the infimum.

Next example presents a situation in which a completely monotone lower probability on a variable is obtained through a probabilistic extension procedure, and how this enables the use of Choquet integral.

**Example 2.** Consider the situation of Example 1. After 5 years of monitoring it has been observed that the internal temperature never exceeds  $10^\circ\text{C}$  during 24:00–6:00, it is always comprised between  $5^\circ\text{C}$  and  $15^\circ\text{C}$  during 6:00–12:00 and 18:00–24:00, while it never decreases below  $5^\circ\text{C}$  during 12:00–18:00. Nevertheless, no relevant connection with relative humidity has been observed due to the variability of the plants stored in the refrigerating room.

Thus denote with  $Z$  the variable expressing the part of the day ranging in  $\mathcal{C}_Z = \{D_1, D_2, D_3, D_4\}$

with  $D_1 = "24:00-6:00"$ ,  $D_2 = "6:00-12:00"$ ,  $D_3 = "12:00-18:00"$  and  $D_4 = "18:00-24:00"$ .

We are interested in deriving a set of probabilistic fuzzy IF-THEN rules related to the vector  $(Y, Z)$ . To simplify the problem, we discretize the range of  $X$  putting  $A_1 = [0, 5]$  and  $A_i = (5(i-1), 5i]$ ,  $i = 2, 3, 4$ , and the range of  $Y$  putting  $B_1 = [0, \frac{1}{5}]$  and  $B_j = (\frac{j-1}{5}, \frac{j}{5}]$ ,  $j = 2, \dots, 5$  thus the joint distribution on  $(X, Y)$  becomes, for  $i = 1, \dots, 4$  and  $j = 1, \dots, 5$ ,

$$P(A_i \wedge B_j) = \frac{250(i^3 - (i-1)^3) + 3(j^2 - (j-1)^2)}{80300}.$$

Moreover, also the membership functions are aggregated as

$$P(E_{\varphi_h} | A_i) = \frac{\int_{A_i} P(E_{\varphi_h} | x) f_X(x) dx}{\int_{A_i} f_X(x) dx},$$

$$P(E_{\psi_k} | B_j) = \frac{\int_{B_j} P(E_{\psi_k} | y) f_Y(y) dy}{\int_{B_j} f_Y(y) dy}.$$

Thus we have

$$\mu_{\varphi_1}(A_i) = P(E_{\varphi_1} | A_i) = \begin{cases} 0.7424, & \text{if } i = 1, \\ 0.0741, & \text{if } i = 2, \\ 0, & \text{otherwise,} \end{cases}$$

$$\mu_{\varphi_2}(A_i) = P(E_{\varphi_2} | A_i) = \begin{cases} 0.2576, & \text{if } i = 1, \\ 0.9259, & \text{if } i = 2, \\ 0.8414, & \text{if } i = 3, \\ 0.1018, & \text{if } i = 4, \end{cases}$$

$$\mu_{\varphi_3}(A_i) = P(E_{\varphi_3} | A_i) = \begin{cases} 0.1586, & \text{if } i = 3, \\ 0.8982, & \text{if } i = 4, \\ 0, & \text{otherwise,} \end{cases}$$

and analogously

$$\mu_{\psi_1}(B_j) = P(E_{\psi_1} | B_j) = \begin{cases} 1, & \text{if } j=1, \\ 0.7499, & \text{if } j=2, \\ 0, & \text{otherwise,} \end{cases}$$

$$\mu_{\psi_2}(B_j) = P(E_{\psi_2} | B_j) = \begin{cases} 0, & \text{if } j=1, \\ 0.2501, & \text{if } j=2, \\ 1, & \text{otherwise.} \end{cases}$$

If we extend the above globally coherent assessment  $P(\cdot | \cdot)$  on the Boolean algebra generated by  $\{D_i, B_j\}_{i=1, \dots, 4, j=1, \dots, 5}$ , the corresponding lower probability  $\underline{P}$  is completely monotone (i.e., it is a belief function) [25].

Related to the variable  $Z$  we consider the properties  $\nu_1 = "light"$  and  $\nu_2 = "dark"$  corresponding to the fuzzy subsets  $E_{\nu_1}^*$  and  $E_{\nu_2}^*$ , having membership

$$\mu_{\nu_1}(D_i) = \begin{cases} 0, & \text{if } i = 1, \\ \frac{1}{2}, & \text{if } i = 2, 4, \\ 1, & \text{if } i = 3, \end{cases}$$

$$\mu_{\nu_2}(D_i) = \begin{cases} 1, & \text{if } i = 1, \\ \frac{1}{2}, & \text{if } i = 3, 4, \\ 0, & \text{if } i = 2. \end{cases}$$

In this context we take the minimum  $t$ -norm  $T_M$  (which is a coherent choice, see Appendix A). In

order to compute the coherent lower  $T_M$ -extension on  $E_{\nu_i} | E_{\psi_j}$  we notice that, for  $i, j = 1, 2$ ,

$$\underline{P}_{T_M}(E_{\nu_i} | E_{\psi_j}) = \frac{\underline{P}_{T_M}(E_{\nu_i} \wedge E_{\psi_j})}{\underline{P}_{T_M}(E_{\psi_j})},$$

i.e., the denominators are the precise probabilities  $P_{T_M}(E_{\psi_1}) = 0.3492$  and  $P_{T_M}(E_{\psi_2}) = 0.6508$ , thus it is sufficient to compute the lower probability  $\underline{P}_{T_M}(E_{\nu_i} \wedge E_{\psi_j})$  for which, being  $\underline{P}$  completely monotone, we can use the Choquet integral.

For example, one has  $\underline{P}_{T_M}(E_{\nu_1} \wedge E_{\psi_1}) = 0.2609$  that corresponds to the imprecise probabilistic fuzzy IF-THEN rule:

- IF  $Y$  IS dry THEN  
–  $Z$  IS light,  $\underline{P}_{T_M}(E_{\nu_1} | E_{\psi_1}) = 0.7471$ .

We stress that the Choquet integral cannot be used to compute  $\underline{P}_{T_M}(E_{\psi_j} | E_{\nu_i})$ , but the general extension procedure for coherent conditional lower probabilities must be applied, instead (see [9, 14]).

In previous example, since the fuzzy information is related to a partition finer than that where the probability is assessed, we need to aggregate the membership functions.

Example 3 considers "certain" fuzzy IF-THEN rules (i.e., for which the corresponding probability is 1) and presents a relevant notion of *entailment*, consisting in the value 1 as unique coherent extension for the new fuzzy conditional event  $E_{\psi} | E_{\varphi}$ .

**Example 3.** The headmaster of a primary school is interested in monitoring the obesity among the pupils of the school, whose age is between 5 and 10 years. The Body Mass Index (BMI) (see for instance [6]) computed as

$$BMI = \text{mass}/(\text{height})^2,$$

is a known statistical measure used for a qualitative (crisp) evaluation of the physique by means of a linguistic label ( $\varphi_1 = "sever underweight"$ ,  $\varphi_2 = "underweight"$ ,  $\varphi_3 = "normal"$ ,  $\varphi_4 = "overweight"$ ,  $\varphi_5 = "obese"$ ).

Thus consider the variables  $X$  and  $Y$  expressing, respectively, the mass in Kg and the height in m, with ranges (due the particular age of the pupils)  $C_X = (0, 70]$  and  $C_Y = (0, 1.5]$ .

The BMI does not take into account the age of the person neither other environmental features so, due to the particular population under study, the linguistic labels  $\varphi_i$  are fuzzified introducing the fuzzy subsets  $E_{\varphi_i}^*$ ,  $i = 1, \dots, 5$ , related to the BMI (that we denote by means of the variable  $Z$ ).

On the other hand, on the variable  $X$  the linguistic properties  $\psi_1 = "very short"$ ,  $\psi_2 = "short"$ ,  $\psi_3 = "medium"$ ,  $\psi_4 = "tall"$ ,  $\psi_5 = "very tall"$  are considered, with corresponding fuzzy subsets of  $C_X$ ,  $E_{\psi_j}^*$ ,  $j = 1, \dots, 5$ . Analogously, on the variable  $Y$  the linguistic properties  $\nu_1 = "very thin"$ ,  $\nu_2 =$

“thin”,  $\psi_3$  = “normal”,  $\psi_4$  = “fat”,  $\psi_5$  = “very fat” are considered, with corresponding fuzzy subsets of  $\mathcal{C}_Y$ ,  $E_{\nu_h}^*$ ,  $h = 1, \dots, 5$ .

Suppose that the assessment  $P$  on

$$\{E_{\varphi_5}, E_{\psi_1} \wedge E_{\psi_2}, E_{\nu_1} \wedge E_{\nu_2}, E_{\varphi_5} \wedge E_{\psi_j} \wedge E_{\nu_h}\}_{j,h=1,2}$$

(that can be computed as the coherent  $T$ -extension of a joint probability on  $(X, Y)$  together with the membership functions) is everywhere positive with  $P(E_{\varphi_5}) = \alpha$ ,  $P(E_{\varphi_5} \wedge E_{\psi_2} \wedge E_{\nu_2}) = \beta$  and  $\alpha > \beta$ .

Then the headmaster finds “natural” to consider the following “certain” fuzzy IF-THEN rules:

- IF  $X$  IS short AND  $Y$  IS fat THEN
  - $Z$  IS obese,  $P(E_{\varphi_5}|E_{\psi_2} \wedge E_{\nu_2}) = 1$ ;
- IF  $X$  IS very short THEN
  - $X$  IS short,  $P(E_{\psi_2}|E_{\psi_1}) = 1$ ;
- IF  $Y$  IS very fat THEN
  - $Y$  IS fat,  $P(E_{\nu_2}|E_{\nu_1}) = 1$ .

Trivial computations show that the value 1 is the unique coherent extension of the above assessment on the events  $\{E_{\varphi_5}|E_{\psi_j} \wedge E_{\nu_h}\}_{j,h=1,2}$ . Thus, one entails the further “certain” fuzzy IF-THEN rules:

- IF  $X$  IS very short AND  $Y$  IS fat THEN
  - $Z$  IS obese,  $P(E_{\varphi_5}|E_{\psi_1} \wedge E_{\nu_2}) = 1$ ;
- IF  $X$  IS short AND  $Y$  IS very fat THEN
  - $Z$  IS obese,  $P(E_{\varphi_5}|E_{\psi_2} \wedge E_{\nu_1}) = 1$ ;
- IF  $X$  IS very short AND  $Y$  IS very fat THEN
  - $Z$  IS obese,  $P(E_{\varphi_5}|E_{\psi_1} \wedge E_{\nu_1}) = 1$ .

On the contrary, the extension on  $E_{\psi_2} \wedge E_{\nu_2}|E_{\varphi_5}$  is unique but strictly less than 1, thus we get the probabilistic fuzzy IF-THEN rule

- IF  $Z$  IS obese THEN
  - $X$  IS short AND  $Y$  IS fat,  
 $P(E_{\psi_2} \wedge E_{\nu_2}|E_{\varphi_5}) = \frac{\beta}{\alpha}$ .

## A. Appendix

A real function  $P$  defined on an arbitrary set of conditional events  $\mathcal{G} = \{E_i|H_i\}_{i \in I}$  is a *coherent conditional probability (assessment)* if and only if is the restriction of a *conditional probability* [18, 19, 24] defined on  $\mathcal{B} \times \mathcal{H} \supseteq \mathcal{G}$ , where  $\mathcal{B}$  is a Boolean algebra and  $\mathcal{H} \subseteq (\mathcal{B} \setminus \{\emptyset\})$  is an additive set (i.e., it is closed under finite logical sums).

For coherent assessments the following fundamental result, essentially due to de Finetti [19], holds:

**Theorem 1.** *Let  $\mathcal{G}, \mathcal{G}'$  be arbitrary sets of conditional events with  $\mathcal{G} \subset \mathcal{G}'$ ,  $P$  a real function on  $\mathcal{G}$ . There exists a coherent conditional probability  $P'$  extending  $P$  on  $\mathcal{G}'$  if and only if  $P$  is a coherent conditional probability on  $\mathcal{G}$ .*

When  $\mathcal{G}' = \mathcal{G} \cup \{E|H\}$ , then the coherent values for the probability of  $E|H$  are a closed interval  $[p_*, p^*]$ . Given a coherent conditional probability  $P$  on  $\mathcal{G} = \{E_i|H_i\}_{i \in I}$  its extension to  $\mathcal{G}' = \mathcal{G} \cup \{E_j|H_j\}_{j=1, \dots, m}$  requires to proceed event-wise, choosing at each step a value in the corresponding extension interval. Previous discussion highlights that the final coherent conditional probability  $P'$  on  $\mathcal{G}'$  depends on the chosen order for the extension on the  $E_j|H_j$ ’s and on the chosen value inside each extension interval: thus, any choice could lead to a loss of information. Hence, in order to preserve all the available information, any choice is avoided and so the extension is carried on the  $E_j|H_j$ ’s separately, obtaining in this way a family of extension intervals  $[p_{j*}, p_j^*]$ ’s. In other words, this approach to extension generates a class

$$\mathbf{P} = \{P^\gamma : \mathcal{G}' \rightarrow [0, 1]\}_{\gamma \in \Gamma}$$

of coherent conditional probabilities on  $\mathcal{G}'$  extending  $P$ . The entire information of the class  $\mathbf{P}$  is summarized by the corresponding lower and upper envelopes  $\underline{P}$  and  $\overline{P}$  on  $\mathcal{G}'$ , that are *coherent conditional lower and upper probabilities* [7]. Recall that the functions  $\underline{P}$  and  $\overline{P}$  are *conjugated* in the sense that, if  $E|H, E^c|H \in \mathcal{G}'$ , then it must hold

$$\underline{P}(E|H) = 1 - \overline{P}(E^c|H). \quad (1)$$

Notice that the functions  $\underline{P}$  and  $\overline{P}$  are conditional assessments satisfying weaker properties w.r.t. a coherent conditional probability.

### A.1. Fuzzy sets as coherent conditional probabilities

Let  $X$  be a (non-necessarily numerical) variable, with range  $\mathcal{C}_X$ , and, for any  $x \in \mathcal{C}_X$ , let us indicate by  $x$  the event  $(X = x)$ .

Let  $\varphi$  be any *property* related to the variable  $X$  and let us refer to the state of information of a real (or fictitious) person that will be denoted by “You”. Moreover, consider the Boolean event

$$E_\varphi = \text{“You claim that } X \text{ is } \varphi\text{”}.$$

A coherent conditional probability assessment  $\{P(E_\varphi|x)\}_{x \in \mathcal{C}_X}$  measures the degree of belief of You in  $E_\varphi$ , when  $X$  assumes the different values of its range. We stress the total freedom of choice of the function  $P(E_\varphi|\cdot)$  in fact, by Corollary 1 in [14], any assessment  $\{P(E_\varphi|x)\}_{x \in \mathcal{C}_X}$  satisfying the trivial consistency condition

$$(L1) \quad P(E_\varphi|x) = 0, \text{ if } E_\varphi \wedge x = \emptyset \text{ and } P(E_\varphi|x) = 1, \text{ if } x \subseteq E_\varphi, \text{ for } x \in \mathcal{C}_X,$$

(and assuming any value in  $[0, 1]$  otherwise) is a coherent conditional probability assessment.

Then  $P(E_\varphi|\cdot)$  comes out to be a natural interpretation of the membership function  $\mu_\varphi(\cdot)$ , according to [8] (see also [7, 9]).

**Definition 1.** For any variable  $X$  with range  $\mathcal{C}_X$  and a related property  $\varphi$ , a **fuzzy subset**  $E_\varphi^*$  of  $\mathcal{C}_X$  is any pair

$$E_\varphi^* = (E_\varphi, \mu_\varphi),$$

with  $\mu_\varphi(x) = P(E_\varphi|x)$  for every  $x \in \mathcal{C}_X$ .

**Remark 1.** The lower envelope of a class  $\{\mu_\varphi^j\}_{j \in J}$  of membership functions concerning the same property  $\varphi$  w.r.t. the same variable  $X$ , satisfies condition **(L1)** and so is still a coherent conditional probability.

## A.2. Operations

In what follows we consider Frank's  $t$ -norms [26], i.e., those defined for  $x, y \in [0, 1]$  and  $\lambda \in [0, +\infty]$ , as

$$T_\lambda^F(x, y) = \log_\lambda \left( 1 + \frac{(\lambda^x - 1)(\lambda^y - 1)}{\lambda - 1} \right). \quad (2)$$

Distinguished instances of previous class are the minimum  $T_M(x, y) = T_0^F(x, y) = \min\{x, y\}$ , the algebraic product  $T_P(x, y) = T_1^F(x, y) = x \cdot y$  and the Łukasiewicz  $t$ -norm  $T_L(x, y) = T_{+\infty}^F(x, y) = \max\{x + y - 1, 0\}$ .

By referring to [8] we recall the operations between fuzzy subsets: the binary operations of union and intersection and that of complementation can be obtained directly by using the rules of coherent conditional probability and the logical independence between  $E_\varphi$  and  $E_\psi$  w.r.t. the partition  $\{(X = x)\}_{x \in \mathcal{C}_X}$  (or w.r.t.  $X$  for short).

Recall that events  $\{E_i\}_{i=1, \dots, n}$  are *logically independent w.r.t. a partition*  $\{H_j\}_{j \in J}$  whenever for any  $j \in J$ ,  $E_i \wedge H_j \neq \emptyset$  for  $i = 1, \dots, n$  implies  $\bigwedge_{i=1}^n E_i \wedge H_j \neq \emptyset$ . A relevant examples of logical independent events are  $\{E_\varphi, E_{\neg\varphi}\}$  and  $\{E_\varphi, E_\psi\}$  where  $\psi$  is “very  $\varphi$ ”. In fact, we can claim both “ $X$  is  $\varphi$ ” and “ $X$  is  $\neg\varphi$ ”, or only one of them or finally neither of them (similarly for  $\varphi$  and “very  $\varphi$ ”). As proved in [8], for any given  $x$  in the range of  $X$ , the assessment  $P(E_\varphi \wedge E_\psi|x)$  is coherent if and only if it holds

$$\begin{aligned} T_L(P(E_\varphi|x), P(E_\psi|x)) &\leq P(E_\varphi \wedge E_\psi|x) \\ &\leq T_M(P(E_\varphi|x), P(E_\psi|x)), \end{aligned} \quad (3)$$

and all the values in the above interval are coherent when  $E_\varphi$  and  $E_\psi$  are logically independent w.r.t.  $X$ .

Given two fuzzy subsets  $E_\varphi^*$ ,  $E_\psi^*$ , corresponding to the same variable  $X$ , with the events  $E_\varphi$ ,  $E_\psi$  logically independent w.r.t.  $X$  let us define:

$$E_{\varphi \wedge \psi}^* = (E_\varphi \wedge E_\psi, \mu_{\varphi \wedge \psi}); E_{\varphi \vee \psi}^* = (E_\varphi \vee E_\psi, \mu_{\varphi \vee \psi})$$

with  $\mu_{\varphi \wedge \psi}(x) = P(E_\varphi \wedge E_\psi|x)$ ,  $\mu_{\varphi \vee \psi}(x) = P(E_\varphi \vee E_\psi|x)$ . Since from probability rules

$$P(E_\varphi \vee E_\psi|x) = P(E_\varphi|x) + P(E_\psi|x) - P(E_\varphi \wedge E_\psi|x).$$

then we can compute  $\mu_{\varphi \wedge \psi}, \mu_{\varphi \vee \psi}$  through any  $t$ -norm and its dual  $t$ -conorm of Frank class (whenever  $E_\varphi$ ,  $E_\psi$  logically independent w.r.t.  $X$ ).

Notice that the hypothesis of logical independence w.r.t.  $X$  is essential. For the events  $E_\varphi$ ,  $E_\psi$ , that are not logically independent w.r.t.  $X$ , the value of  $P(E_\varphi \vee E_\psi|x)$  is obtained by a coherent extension: the value satisfies inequality (3), but not all the values in such interval can be assumed. For example, it holds  $P(E_\varphi \wedge E_\psi|x) = P(E_\varphi|x)$  thus, in this case, only the right extreme of (3) is a coherent choice.

We finally recall that in this context the complement of a fuzzy subset is defined as

$$(E_\varphi^*)' = (E_{\neg\varphi}, \mu_{\neg\varphi}) = (E_{\neg\varphi}, 1 - \mu_\varphi).$$

Obviously it holds  $E_{\neg\varphi} \neq (E_\varphi)^c$ , since the propositions “You claim  $X$  is  $\neg\varphi$ ” and “You do not claim  $X$  is  $\varphi$ ” are logically independent. Then, while  $E_\varphi \vee (E_\varphi)^c = \Omega$ , one has  $E_\varphi \vee E_{\neg\varphi} \subseteq \Omega$  and so  $(E_\varphi^*)'$  is generally only a fuzzy subset of the universe  $\mathcal{C}_X$ , for which  $\mu_{\varphi \vee \neg\varphi}(x) = \mu_\varphi(x) + \mu_{\neg\varphi}(x) - \mu_{\varphi \wedge \neg\varphi}(x)$ .

A similar discussion holds even for  $E_\varphi$  and  $E_\psi$ , where  $\psi$  is the superlative of  $\varphi$ .

For two fuzzy subsets  $E_\varphi^*$  and  $E_\psi^*$ , corresponding to the random quantities  $X$  and  $X'$ , respectively, the following choice for the membership of conjunction and disjunction is coherent [9]:

$$\mu_{\varphi \wedge \psi}(x, x') = P(E_\varphi \wedge E_\psi|x \wedge x'),$$

$$\mu_{\varphi \vee \psi}(x, x') = P(E_\varphi \vee E_\psi|x \wedge x'),$$

with, for any  $x \in \mathcal{C}_X$  and  $x' \in \mathcal{C}_{X'}$ , the only constraints:

$$\begin{aligned} T_L(\mu_\varphi(x), \mu_\psi(x')) &\leq \mu_{\varphi \wedge \psi}(x, x') \\ &\leq T_M(\mu_\varphi(x), \mu_\psi(x')) \end{aligned} \quad (4)$$

and

$$\mu_{\varphi \vee \psi}(x, x') = \mu_\varphi(x) + \mu_\psi(x') - \mu_{\varphi \wedge \psi}(x, x'). \quad (5)$$

## A.3. Probability of “fuzzy events”

In this context, the concept of fuzzy event, as introduced by Zadeh, is an ordinary event of the kind

$$E_\varphi = \text{“You claim that } X \text{ is } \varphi\text{”}.$$

For any finitely additive probability measure  $p$  on the Boolean complete algebra generated by  $\mathcal{C}_X$ , the global assessment  $\{\mu_\varphi, p\}$  is coherent [7, 6] (see also [37, 42]) and so coherently extendible to  $E_\varphi$ . According to Theorem 3 in [15], the only coherent value for the probability of  $E_\varphi$  is

$$P(E_\varphi) = \int \mu_\varphi(x) dp(x), \quad (6)$$

which formally coincides with Zadeh’s definition of probability of a “fuzzy event” [48]. Actually it generalizes Zadeh’s definition since  $p$  is a coherent

finitely additive (non-necessarily  $\sigma$ -additive) probability.

Now let  $\mathcal{E}^* = \{E_{\varphi_1}^*, \dots, E_{\varphi_1}^*\}$  be a finite collection of fuzzy subsets

$$E_{\varphi_i}^* = (E_{\varphi_i}, \mu_{\varphi_i})$$

of  $\mathcal{C}_X$ ,  $i = 1, \dots, n$ , related to the (possibly coincident) components  $X_j$  of a  $m$ -dimensional vector  $X$  (with the events  $\{E_{\varphi_i}\}_{i=1, \dots, n}$  logically independent with respect to the partition generated by  $X$ ).

For a given Frank's  $t$ -norm  $T$ , denote  $\langle \mathcal{E}^* \rangle_T$  the closure of  $\mathcal{E}^*$  with respect to fuzzy intersection and union related to  $T$  as defined in previous section. Again the global assessment  $\{\mu_{\varphi_i}, P\}$ , where  $P$  is related to the complete Boolean algebra spanned by the random vector  $(X_1, \dots, X_m)$ , is coherent. Moreover it is easy to prove (see [8]) that for a fixed Frank's  $t$ -norm  $T$  [26], the probability assessment

$$\{P_T(E_{\varphi_i}), P_T(E_{\varphi_i} \wedge E_{\varphi_j})\}$$

is still coherent, with

$$P_T(E_{\varphi_i}) = \int \mu_{\varphi_i}(x_i) dP_i(x_i),$$

$$P_T(E_{\varphi_i} \wedge E_{\varphi_j}) = \int T(\mu_{\varphi_i}(x_i), \mu_{\varphi_j}(x_j)) dP_{ij}(x_{ij})$$

(where  $P_i [P_{ij}]$  is the marginal on  $X_i [(X_i, X_j)]$ ).

So we can extend  $P_T$  to events  $E_{\varphi_i} \vee E_{\varphi_j}$ : this extension is univocally determined by coherence as

$$P_T(E_{\varphi_i} \vee E_{\varphi_j}) = P_T(E_{\varphi_i}) + P_T(E_{\varphi_j}) - P_T(E_{\varphi_i} \wedge E_{\varphi_j}).$$

**Remark 2.** The condition of logical independence of events  $E_{\varphi_i}, E_{\varphi_j}$  with respect to  $X$  is crucial for proving all the above assertions [15]. For instance the same formula cannot be used for obtaining the coherent extension of  $P_T$  to  $E_{\varphi_i} | E_{\varphi_i}$  which is necessarily 1, independently of the Frank's  $t$ -norm used for computing intersection and union between the fuzzy sets related to the family  $\{E_{\varphi_i}\}_{i=1, \dots, n}$ .

The above assessment  $P_T$  is a coherent conditional probability, so by Theorem 1 it can be furthermore extended to any conditional event  $A|B$  where  $A, B$  are events of the Boolean algebra  $\mathcal{B}$  generated by  $\{E_{\varphi_i}\}_{i=1, \dots, n} \cup \{(X = x)\}_{x \in \mathcal{C}_X}$ , with  $B \neq \emptyset$ . This extension is not unique in general, but for the events  $A = E_{\varphi_i}$  and  $B = E_{\varphi_j}$ ,  $i \neq j$ , with  $P_T(E_{\varphi_j}) > 0$  the only coherent extension is:

$$P_T(E_{\varphi_i} | E_{\varphi_j}) = \frac{\int T(\mu_{\varphi_i}(x_i), \mu_{\varphi_j}(x_j)) dP_{ij}(x_{ij})}{\int \mu_{\varphi_j}(x_j) dP_j(x_j)}.$$

We call the above extension of  $P$  a coherent  $T$ -extension.

## References

[1] P. Berti, E. Dreassi, P. Rigo, A consistency theorem for regular conditional distributions, *Stochastics*, to appear, DOI:10.1080/17442508.2011.653644, 2012.

[2] A. Brozzi, A. Capotorti, B. Vantaggi, Incoherence correction strategies in statistical matching. *International Journal of Approximate Reasoning*, vol. 53(8), pp. 1124–1136, 2012.

[3] A. Capotorti, G. Figá-Talamanca, On an implicit assessment of fuzzy volatility in the Black and Scholes environment. *Fuzzy Sets and Systems*, In Press, 2013.

[4] C. Chen, T. Xiao, Probabilistic Fuzzy Control of Mobile Robots for Range Sensor Based Reactive Navigation. *Intelligent Control and Automation*, vol. 2, pp. 77–85, 2011.

[5] G. Coletti, Coherent numerical and ordinal probabilistic assessments, *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 24, pp. 1747–1754, 1994.

[6] G. Coletti, O. Gervasi, S. Tasso, B. Vantaggi, Generalized Bayesian inference in a fuzzy context: From theory to a virtual reality application, *Computational Statistics & Data Analysis*, vol. 56 (4), pp. 967–980, 2012.

[7] G. Coletti, R. Scozzafava, *Probabilistic logic in a coherent setting*. Trends in logic, vol. 15, Kluwer, Dordrecht, 2002.

[8] G. Coletti, R. Scozzafava, Conditional Probability, Fuzzy Sets, and Possibility: a Unifying View, *Fuzzy Sets and Systems*, vol. 144, pp. 227–249, 2004.

[9] G. Coletti, R. Scozzafava, Conditional Probability and Fuzzy Information, *Computational Statistics & Data Analysis*, vol. 51, pp. 115–132, 2006.

[10] G. Coletti, R. Scozzafava, Toward a general theory of conditional beliefs, *International Journal of Intelligent Systems*, vol. 21, pp. 229–259, 2006.

[11] G. Coletti, R. Scozzafava, B. Vantaggi, Coherent Conditional Probability, Fuzzy Inclusion and Default Rules. *Proc. of First World Conference on Soft Computing*, San Francisco 2011.

[12] G. Coletti, R. Scozzafava, B. Vantaggi, Weak implication in terms of conditional uncertainty measures, *Lecture Notes in Computer Science*, vol. 4724, pp. 139–150, 2007.

[13] G. Coletti, B. Vantaggi, Representability of ordinal relations on a set of conditional events. *Theory and Decision*, vol. 60(2-3), pp. 137–174, 2006.

[14] G. Coletti, B. Vantaggi, Probabilistic reasoning with vague information, proceedings of the *2nd World Conference on Soft Computing*, Baku, Azerbaijan, December 3-5, 2012.

[15] G. Coletti, B. Vantaggi, Hybrid models: probabilistic and fuzzy information, *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, Kruse, R.; Berthold, M.R.; Moewes, C.; Gil, M.Á.; Grzegorzewski, P.; Hryniewicz, O. (Eds.), pp. 389–398, 2012.

[16] R. Coppi, M.A. Gil, H.A.L. Kiers (Eds.), The Fuzzy Approach to Statistical Analysis, *Compu-*

- tational Statistics & Data Analysis*, 51, 1–452, 2006.
- [17] M. D’Orazio, M. Di Zio, M. Scanu, *Statistical Matching: Theory and Practice*, Wiley Series in Survey Methodology, John Wiley & Sons, Ltd., Chichester, 2006.
  - [18] B. de Finetti, Sull’impostazione assiomatica del calcolo delle probabilità, *Annali Univ. Trieste*, vol. 19, pp. 3–55, 1949. (Engl. transl.: Ch.5 of *Probability, Induction and Statistics*, Wiley, London, 1972.)
  - [19] B. de Finetti, *Teoria della probabilità*, Einaudi, Torino, 1974. (Engl. transl.: *Theory of Probability*, voll. 1 and 2., Wiley, Chichester, 1970.)
  - [20] D. Denneberg, *Non-Additive Measure and Integral*. Kluwer, Dordrecht, 1994.
  - [21] D. Dubois, S. Moral, H. Prade, A semantics for possibility theory based on likelihoods, *Journal of Mathematical Analysis and Applications* 205, 359–380, 1997.
  - [22] D. Dubois, H. Prade, The Three Semantics of Fuzzy Sets, *Fuzzy Sets and Systems*, 90, 141–150 1997.
  - [23] M.A. Gil, M. Lopez-Diaz, D.A. Ralescu (Eds.), Fuzzy sets and probability/Statistics Theories, *Fuzzy Sets and Systems* 157, 2545–2682, 2006.
  - [24] L.E. Dubins, Finitely Additive Conditional Probabilities, Conglomerability and Disintegration, *The Annals of Probability*, vol. 3, pp. 89–99, 1975.
  - [25] R. Fagin, J. Halpern, Uncertainty, belief and probability, *Computational Intelligence*, vol. 7, pp. 160–173, 1991.
  - [26] M.J. Frank, On the simultaneous associativity of  $F(x, y)$  and  $x + y - F(x, y)$ , *Aequationes Mathematicae*, vol. 19, pp. 194–226, 1979.
  - [27] A. Gilio, Probabilistic Reasoning Under Coherence in System P, *Annals of Mathematics and Artificial Intelligence*, vol. 34, pp. 5 – 34, 2002.
  - [28] A. Gilio, D. Over, The psychology of inferring conditionals from disjunctions: A probabilistic study. *Journal of Mathematical Psychology*, vol. 56(2), pp. 118–131, 2012.
  - [29] A. Gilio, G. Sanfilippo, Probabilistic entailment in the setting of coherence: The role of quasi conjunction and inclusion relation. *International Journal of Approximate Reasoning*, In Press, 2013.
  - [30] D. Lehmann, M. Magidor, What does a conditional knowledge base entail?, *Artificial Intelligence*, vol. 55, pp. 1–60, 1992.
  - [31] Z. Liu and H.-X. Li, A probabilistic fuzzy logic system for modeling and control, *IEEE Transactions on Fuzzy Systems*, vol. 13(6), pp. 848–859, 2005.
  - [32] A.H. Meghdadi, M.-R. Akbarzadeh-T, Probabilistic Fuzzy Logic and Probabilistic Fuzzy Systems, *Proc. IEEE International Fuzzy Systems Conference*, pp. 1127–1130, 2001.
  - [33] E. Miranda, Updating coherent previsions on finite spaces, *Fuzzy Sets and Systems*, 160(9), 1286–1307, 2009.
  - [34] F. Montagna, A Notion of Coherence for Books on Conditional Events in Many-valued Logic, *Journal of Logic and Computation*, 44(3), pp. 563–593, 2012.
  - [35] D. Mundici, Averaging the truth value in Łukasiewicz logic, *Studia Logica*, 55, pp. 113–127, 1995.
  - [36] D. Mundici, Bookmaking over infinite-valued events *International Journal of Approximate Reasoning*, 46, pp. 223–240, 2006.
  - [37] E. Regazzini, Finitely additive conditional probabilities, *Rendiconti del Seminario Matematico e Fisico di Milano*, vol. 55(1), pp. 69–89, 1985.
  - [38] D.B. Rubin, Statistical matching using file concatenation with adjusted weights and multiple imputations, *Journal of Business and Economic Statistics*, 2 (1), 87–94, 1986.
  - [39] K. Rudnik, A. Walaszek-Babiszewska, Probabilistic-fuzzy knowledge-based system for managerial application. *Management and Production Engineering Review*, vol. 3(1), pp. 49–61, 2012.
  - [40] R. Scozzafava, B. Vantaggi, Fuzzy Inclusion and Similarity through Coherent Conditional Probability, *Fuzzy Sets and Systems*, vol. 160, pp. 292–305, 2009.
  - [41] M. Tang, X. Chen, W. Hu, W. Yu, Generation of a probabilistic fuzzy rule base by learning from examples. *Information Sciences*, vol. 217, pp. 21–30, 2012.
  - [42] B. Vantaggi, Statistical matching of multiple sources: a look through coherence. *International Journal of Approximate Reasoning*, vol. 49, pp. 701–711, 2008.
  - [43] R. Viertl, *Statistical methods for non-precise data*. CRC Press, Boca Raton, 1996.
  - [44] R. Viertl, Is it necessary to develop a fuzzy Bayesian inference? in: R. Viertl (Ed.) *Probability and Bayesian Statistics*, 471–475, 1987.
  - [45] P. Walley, *Statistical reasoning with imprecise probabilities*. London: Chapman and Hall, 1991.
  - [46] L. Waltman, U. Kaymak, J. van den Berg, Maximum Likelihood Parameter Estimation in Probabilistic Fuzzy Classifiers, *Proc. IEEE International Conference on Fuzzy Systems*, pp. 1098–1103, 2005.
  - [47] L.A. Zadeh, Fuzzy sets, *Information and Control*, vol. 8, pp. 338–353, 1965.
  - [48] L.A. Zadeh, Probability measures of fuzzy events, *Journal of Mathematical Analysis and Applications*, vol. 23, pp. 421–427, 1968.
  - [49] L.A. Zadeh, Toward a perception-based theory of probabilistic reasoning with imprecise probabilities, *Journal of Statistical Planning and Inference*, vol. 105, pp. 233–264, 2002.