OPEN ACCESS

University of BRISTOL

Baldwin, J. F., Lawry, J., & Martin, T. P. (1996). A Note on Probability /Possibility Consistency for Fuzzy Events. In IPMU 96.

Link to publication record in Explore Bristol Research
PDF-document

A Note on Probability / Possibility Consistency for Fuzzy Events

J.F.Baldwin, J.Lawry and T.P.Martin

A.I. Group, Department of Engineering Mathematics, University of Bristol, Bristol BS8 1TR, United Kingdom

Abstract
It is noted that the standard t-norm definitions of conditional probability of fuzzy events are probability / possibility inconsistent. Two alternative definitions are proposed and their consistency with two definitions of the possibility of fuzzy events is investigated.

1 Introduction

Modelling real world problems typically involves processing uncertainty of two distinct types. These are uncertainty arising from a lack of knowledge relating to concepts which, in the sense of classical logic, may be well defined and uncertainty due to inherent vagueness in concepts themselves. Traditionally the above are modelled in terms of probability theory and fuzzy set theory respectively. Furthermore, there are many situations where we have insufficient information regarding vague or fuzzy concepts. That is where both types of uncertainty are present. This suggests the need for theories of the probability and possibility of fuzzy events. In [10] Zadeh proposes that the probability of fuzzy events of the form X is f , where X is a random variable into some universe and f is a fuzzy subset of , be defined as the expected value of the membership function of f with respect to the probability distribution of X . Zadeh then gives a definition of the conditional probability of fuzzy events which is a special case of what we shall refer to as the t-norm definition of the conditional probability of fuzzy events.

Definition 1.1 (t-norm definition of the conditional probability of fuzzy events)
For f and g fuzzy subsets of
$$Prob\big(\text{X is } f | \text{X is } g\big) =$$

$$\frac{Prob\big(\text{X is } f \,_t\, g\big)}{Prob\big(\text{X is } g\big)}$$ where $_t$ is some t-norm corresponding to Zadeh's definition when $_t$ is the product conjunction.

2 Possibility Distributions and Fuzzy Events

According to Zadeh (see [11]) the fuzzy event X is g is interpreted to mean that the random variable X has the possibility distribution g so that

$$Pos\big(\text{X} \in \text{S}|\text{X is } g\big) = \sup_{x \in S} g(x)$$

Zadeh then extends the above notion of possibility to fuzzy events as follows:

Definition 2.1 (Zadeh)
For f and g fuzzy subsets of
$$Pos_Z\big(\text{X is } f|\text{X is } g\big)$$

$$= \sup_x \min\Big(f(x), g(x)\Big)$$

There are, of course, innumerable alternative extensions of possibility measures to fuzzy sets consistent with the crisp case, however, for the scope of this paper we shall consider only the following:

Definition 2.2
For f and g fuzzy subsets of
$$Pos_A\big(\text{X is } f|\text{X is } g\big)$$

$$= \int_0^1 Pos\big(\text{X} \in f_y\big|\text{X is } g\big) \, dy$$

provided this integral exists and is left undefined otherwise. Here, in accordance with standard notation, f_y denotes the

y'th -cut of f .

In the sequel we shall provide some justification for this alternative definition in terms of a general mechanism for extending set theoretic operations to fuzzy sets.
Smets [9] extends the notion of a belief function to fuzzy sets together with the corresponding notion of a plausibility measure. In this framework both of the above definitions are valid measures of

possibility since provided g is a normalised fuzzy set both are consonant plausibility measures.

It should be noted that in the above definitions we have adopted a somewhat non-standard conditional notation for possibility measures defined according to some prior possibility distribution g. It seems natural, though, to regard such a measure of possibility as conditional on the fuzzy event X is g since according to Zadeh (see [11]) to assert that such an event occurs is to assert precisely that X has the possibility distribution g.

3 The Principle of Probability / Possibility Consistency

Now given the above interpretation of fuzzy events it would seem a requirement of any definition of the conditional probability of fuzzy events that it should be probability / possibility consistent in the following sense;

for any f and g fuzzy subsets of

$$Prob\big(X \text{ is } f | X \text{ is } g\big)$$

$$Pos\big(X \text{ is } f | X \text{ is } g\big)$$

The notion of probability / possibility consistency was introduced by Zadeh in [11] although it is clear that he had a some what weaker constraint in mind. For example, he comments: "It should be understood, of course, that the possibility / probability consistency principle is not a precise law or a relationship that is intrinsic in the concepts of possibility and probability. Rather it is an approximate formalization of the heuristic observation that a lessening of the possibility of an event tends to lessen its probability - but not vice-versa." An investigation into alternative notions of probability / possibility consistency can be found in Delgado and Moral [5]. For this paper, however, we shall adopt the strong principle as stated above in accordance with Dubois and Prade (see [6])

Of course the principle of probability / possibility consistency has different implications depending on how we extend possibility measures to fuzzy events , however, the following example shows that no t-norm definition of the conditional

probability of fuzzy events satisfies this principle even for the case when f is crisp.

Example

Let X be a random variable into $= \{a, b, c\}$ distributed according to the uniform distribution. Further let $S = \{b, c\}$ and $g = a/1 + b/0.4 + c/0.34$ then

according to the t-norm definition of conditional probability we have that

$$Prob\big(X \quad S | X \text{ is } g\big) =$$

$$\frac{t(1, 0.4) + t(1, 0.35)}{1 + 0.4 + 0.35}$$

Now it is a property of t-norms that $t(x, 1) = t(1, x) = x$ and therefore

$$Prob\big(X \quad S | X \text{ is } g\big) = \frac{0.75}{1.75} \quad 0.4285$$

However

$$Pos\big(X \quad S | X \text{ is } g\big) =$$
$$max\big(0.4, 0.35\big) = 0.4$$

4 Alternative Definitions for the Conditional Probability of Fuzzy Events

In the sequel we propose two alternative definitions for the conditional probability of fuzzy events one being consistent with Pos_Z and the other with Pos_A. Dubois and Prade (see [7]) have proposed a method for extending crisp set theoretic operations to fuzzy sets utilising the notion of -cuts. More specifically if $F{:}2 \quad \mathbf{R}$ then we extend F to fuzzy sets such that for g a fuzzy subset of we have

$$F\big(g\big) = \int_0^1 F\big(g_y\big) \, dy$$

Now a possible justification for this definition is as follows. Consider an intelligent agent faced with the problem of calculating the value of some operation on a fuzzy set where that operation has only been defined for crisp sets. One solution to this problem is for the agent to generate a crisp set from the fuzzy set and to apply the operator to this crisp set. Now a possible mechanism by which the agent could generate a corresponding crisp set is

2

simply to select a value y, a threshold if you like, at random from $[0,1]$ according to the uniform distribution on $[0,1]$ and to take as his crisp set to be the set of all elements in the domain with membership in the fuzzy set greater than or equal to y.

That is he selects the y 'th α-cut of the fuzzy set. Dubois and Prade's definition then corresponds to the expected response of an agent reasoning according to such a non-deterministic mechanism.
Now it is well known that Zadeh's definition of the probability of fuzzy events can de viewed as the extension of a probability measure to fuzzy sets according to the above mechanism (see [6] and [8]). In addition, it is interesting to note that $\boldsymbol{Pos}_A$ can viewed as an extension of possibility measures to fuzzy sets according to the mechanism discussed above. Indeed, we might consider this interpretation as providing some justification for $\boldsymbol{Pos}_A$ as a measure of possibility for fuzzy events.
Suppose then that we apply this method to conditional probability. Of course, here the operator is binary so that in effect our agent must pick two thresholds one for each fuzzy set and in this case we must consider what should be the relationship between the two thresholds. In the following we propose two possible relationships between the thresholds each giving us an alternative definition for the conditional probability of fuzzy events. One possibility is that our agent independently selects two thresholds from $[0,1]$ at random according to the uniform distribution on $[0,1]^2$. This would appear to some extent justifiable since because for each fuzzy set the threshold is chosen entirely at random from $[0,1]$ and is not dependent on the interpretation of the fuzzy set it would seem unreasonable for that threshold to be in any way dependent on some other fuzzy set. This assumption of independence between thresholds yields the following definition of conditional probability of fuzzy events.

Definition 4.1

Let X be a random variable into Ω with probability distribution $\boldsymbol{W}$ then for $\boldsymbol{f}$ and $\boldsymbol{g}$ fuzzy subsets of Ω

$$\boldsymbol{Prob}_{\boldsymbol{W}}^{i}(\text{X is } \boldsymbol{f} \mid \text{X is } \boldsymbol{g}) =$$

$$\int_0^1 \int_0^1 \frac{W(f_y \cap g_s)}{W(g_s)} \, ds \, dy \quad \text{provided this}$$

integral exists and is undefined otherwise.
Now consider $\boldsymbol{F}{:}[0,1]^2 \to [0,1]$ such that

$$F(y,s) = \frac{W(f_y \cap g_s)}{W(g_s)} \quad \text{then clearly}$$

$\boldsymbol{Prob}_{\boldsymbol{W}}^{i}(\text{X is } \boldsymbol{f} \mid \text{X is } \boldsymbol{g})$ is defined if and only if $\boldsymbol{F}$ is lebesgue integrable on $[0,1]^2$. In particular, this does not hold if $\exists \varepsilon > 0$ such that $\forall s \in [1-\varepsilon, 1+\varepsilon]$ $W(g_s) = 0$. A consequence of this is that if $\sup_{x} \mu_g(x) < 1$ then

$\boldsymbol{Prob}_{\boldsymbol{W}}^{i}(\text{X is } \boldsymbol{f} | \text{X is } \boldsymbol{g})$ is undefined for all probability distributions $\boldsymbol{W}$.
Note that the superscript i in the above is used to denote independent threshold.

This definition forms the basis for semantic unification in Fril (see [1], [2] and [3]) which is a logic programming style language with the capability of manipulating both probabilistic and fuzzy uncertainty. It can also be formulated in terms of mass assignments and has a number of desirable properties (see [4] for details). Now, as is shown by the following result, this definition of conditional probability turns out to be consistent with $\boldsymbol{Pos}_A$.

Theorem 4.2
Let $\boldsymbol{f}$ and $\boldsymbol{g}$ be fuzzy subsets of Ω and $\boldsymbol{W}$ be a probability distribution on Ω such that both $\boldsymbol{Prob}_{\boldsymbol{W}}^{i}\left(\text{X is } \boldsymbol{f} | \text{X is } \boldsymbol{g}\right)$ and $\boldsymbol{Pos}_A\left(\text{X is } \boldsymbol{f} | \text{X is } \boldsymbol{g}\right)$ are defined then

$$\boldsymbol{Prob}_{\boldsymbol{W}}^{i}\left(\text{X is } \boldsymbol{f} | \text{X is } \boldsymbol{g}\right)$$
$$\leq \boldsymbol{Pos}_A\left(\text{X is } \boldsymbol{f} | \text{X is } \boldsymbol{g}\right)$$

Proof

Notice that for any $T \subseteq \Omega$ if $s > \sup_{x \in T} g(x)$ then $T \cap g_S = \emptyset$ since if $x \in T \cap g_S$ then $g(x) \geq s$ which is a contradiction. Therefore,

$$\boldsymbol{Prob}^i_{\boldsymbol{W}}(X \text{ is } \boldsymbol{f} \mid X \text{ is } \boldsymbol{g}) =$$

$$\int_0^1 \int_0^1 \frac{W\left(f_y \cap g_s\right)}{W\left(g_s\right)} \, ds\, dy =$$

$$\int_0^1 \int_0^{\sup_x f_y \wedge g(x)} \frac{W\left(f_y \cap g_s\right)}{W\left(g_s\right)} \, ds\, dy$$

$$\int_0^1 \int_0^{\sup_x f_y \wedge g(x)} 1 \, ds\, dy = \int_0^1 \boldsymbol{Pos}\left(X \in f_y \mid X \text{ is } \boldsymbol{g}\right) dy$$

❑

However, as illustrated by the following example, $\boldsymbol{Prob}^i$ is not consistent with $\boldsymbol{Pos}_Z$.

Example 4.3
Let $\boldsymbol{f} = a/1 + c/0.4 + b/0.3$ and $\boldsymbol{g} = b/1 + c/0.95 + a/0.35$ be fuzzy subsets of $\Omega = \{a, b, c\}$ and let X be a random variable into Ω distributed according W such that $W(\{a\}) = 0.989, W(\{b\}) = 0.001$ and $, W(\{c\}) = 0.01$ then

$$\boldsymbol{Prob}^i_{\boldsymbol{W}}\left(X \text{ is } \boldsymbol{f} \mid X \text{ is } \boldsymbol{g}\right) =$$

$$(0.3)(0.35)\,W\big(\{a, c, b\} \mid \{b, c, a\}\big) +$$

$$(0.3)(0.6)\,W\big(\{a, c, b\} \mid \{b, c\}\big) +$$

$$(0.3)(0.05)\,W\big(\{a, c, b\} \mid \{b\}\big) +$$

$$(0.1)(0.35)\,W\big(\{a, c\} \mid \{b, c, a\}\big) +$$

$$(0.1)(0.6)\,W\big(\{a, c\} \mid \{b, c\}\big) +$$

$$(0.6)(0.3)\,W\big(\{a\} \mid \{b, c, a\}\big) \quad 0.105(1) +$$

$$0.18(1) + 0.015(1) + 0.035(0.999) +$$

$$0.06(0.9091) + 0.18(0.989) = 0.579$$

where as

$$\boldsymbol{Pos}_Z\big(X \text{ is } f \mid X \text{ is } \boldsymbol{g}\big) =$$

$$\boldsymbol{max}\left(0.3, 0.35, 0.4\right) = 0.4$$

In Fril the underlying mechanism for manipulating uncertainty is the support

logic calculus (see [1] and [3]) by which a support pair for any statement can be inferred from a knowledge base consisting of facts and rules with associated supports. The method of semantic unification (see [2] and [3]) calculates a support pair for conditional rules of the form $\left(X \text{ is } \boldsymbol{f} \text{ if } X \text{ is } \boldsymbol{g}\right)$ where the upper support corresponds to $\boldsymbol{Pos}_A\left(X \text{ is } f \mid X \text{ is } \boldsymbol{g}\right)$ and the lower support to

$$\boldsymbol{Nec}_A\left(X \text{ is } f \mid X \text{ is } \boldsymbol{g}\right) =$$

$$1 - \boldsymbol{Pos}_A\left(X \text{ is } \bar{f} \mid X \text{ is } \boldsymbol{g}\right)$$

If a point value support is required Fril calculates $\boldsymbol{Prob}^c_{\boldsymbol{W}}\left(X \text{ is } f \mid X \text{ is } \boldsymbol{g}\right)$ where as a default W is assumed to be the uniform distribution.

Now suppose that instead of selecting the threshold for the two fuzzy sets independently our agent selects the same threshold for each so that in effect he selects a single threshold at random from $[0,1]$ according to the uniform distribution on $[0,1]$ and generates both crisp sets relative to this threshold. The motivation for this might be that the agent should select a constant threshold representing his degree of optimism for all fuzzy sets to be acted upon by the operator. The notion of a constant threshold has been discussed in the context of the voting model for fuzzy sets by Baldwin (see [3]) who notes the relationship between this assumption and the choice of min as the conjunction operator for fuzzy sets. According to this constant threshold relationship we have the following definition for conditional probability:

Definition 4.4
Let X be a random variable into Ω with probability distribution W then for $\boldsymbol{f}$ and $\boldsymbol{g}$ fuzzy subsets of Ω

$$\boldsymbol{Prob}^c_{\boldsymbol{W}}(X \text{ is } \boldsymbol{f} \mid X \text{ is } \boldsymbol{g})$$

$$= \int_0^1 \frac{W\left(f_y \cap g_y\right)}{W\left(g_y\right)} \, dy \text{ provided this integral}$$

exists and is undefined otherwise. Now consider $\boldsymbol{F}:[0,1] \to [0,1]$ such that

4

$$F(y) = \frac{W(f_y \cap g_y)}{W(g_y)} \quad \text{then clearly}$$

$Prob_W^c(X$ is f | X is $g)$ is defined if and only if F is lebesgue integrable on $[0,1]$. As with the previous definition this does not hold if $\epsilon > 0$ such that

$y \in [1-\epsilon, 1+\epsilon] \Rightarrow W(g_y) = 0$ so that

$Prob_W^c(X$ is f | X is $g)$ is also undefined for all W if $\sup_x g(x) < 1$.

Note that the superscript c in this definition is used to denote constant threshold.

In this case, as is shown by the following result, we have consistency with Pos_Z

Theorem 4.5
For any fuzzy sets f and g and distribution W such that $Prob_W^c(X$ is $f|X$ is $g)$ is defined we have that
$Prob_W^c(X$ is $f|X$ is $g)$
$\leq Pos_Z(X$ is $f|X$ is $g)$

Proof

Now $Prob_W^c(X$ is $f|X$ is $g)$

$$= \int_0^1 \frac{W(f_y \cap g_y)}{W(g_y)} \, dy$$

$$= \int_0^{\substack{supmin_x(f(x), g(x))}} \frac{W(f_y \cap g_y)}{W(g_y)} \, dy$$

since $f_y \cap g_y =$

$\{x \mid min(f(x), g(x)) \geq y\}$

so that if $y > \sup min_x(f(x), g(x))$
then $f_y \cap g_y = \emptyset$

$\leq Prob_W^c(X$ is $f|X$ is $g)$

$$\int_0^{\substack{supmin_x(f(x), g(x))}} dy$$

$$= \sup_x min(f(x), g(x))$$

so that $\sup_x min(f(x), g(x))$ is an upper bound of $Prob_W^c(X$ is $f|X$ is $g)$

The following example shows, however, that $Prob^c$ is not consistent with Pos_A

Example 4.6
Let $f = a/1 + c/0.4 + b/0.3$ and $g = d/1 + e/0.2 + c/0.4$ be fuzzy subsets of $\Omega = \{a,b,c,d,e\}$ and X be a random variable into Ω then
$Pos_A(X$ is $f|X$ is $g) =$
$0.3 \sup_{x \in \{a,c,b\}} g(x) +$
$0.1 \sup_{x \in \{a,c\}} g(x) = 0.3(0.4) +$
$0.1(0.4) = 0.16$
Now let W be such that $W(\{c\}) = 0.99$ and $W(\{d\}) = 0.01$ then
$Prob_W^c(X$ is $f|X$ is $g) =$
$0.2 W(\{a,c,b\}|\{d,c,e\}) +$
$0.1 W(\{a,c,b\}|\{d,c\}) +$
$0.1 W(\{a,c\}|\{d,c\}) = 0.2(0.99) +$
$0.1(0.99) + 0.1(0.99) = 0.396$

5 Conclusion

The method of extending set theoretic operations to fuzzy sets proposed in [7] has been used to generate two possible definitions for the conditional probability of fuzzy events each being a consequence of the relationship between thresholds chosen for each fuzzy set. The assumption of independent thresholds provides a definition which is consistent with our alternative notion of possibility for fuzzy events where as if both thresholds are assumed to be equal then we have a definition consistent with Zadeh's notion (see [11]). It is, of course, a trivial consequence of these consistency results

that both definitions are probability / possibility consistent for crisp events suggesting that they are more justifiable than the widely accepted t-norm definitions.

6 References

[1] J.F.Baldwin,T.P Martin, B.W Pilsworth, (1988), FRIL Manual (Version 4.0) FRIL systems Ltd, Bristol Business Centre, Maggs House, Queens Road, Bristol, BS8 1QX , U.K

[2]J.F.Baldwin (1990) Computational Models of Uncertainty Reasoning in Expert Systems, Computers Math. Applic. Vol 19 No.11 pp105-119

[3] J.F.Baldwin, T.P Martin, B.W Pilsworth, (1995) FRIL-Fuzzy and Evidential Reasoning in A.I, Research Studies Press, John Wiley

[4] J.F. Baldwin, J. Lawry, T.P. Martin, (1995), A Mass Assignment Theory of the Probability of Fuzzy Events, to appear in Fuzzy Sets and Systems

[5] M.Delgado, S.Moral, (1987) On the Concept of Possibility-Probability Consistency, Fuzzy Sets and Systems 21 pp311-318

[6] D.Dubois , H.Prade, (1990), Scalar Evaluation of Fuzzy Sets: Overview and Applications, Appl. Math. Lett Vol 3, No 2 pp 37-42

[7] D. Dubois, H. Prade, (1990), Measuring Properties of Fuzzy Sets: A General Technique and its use in Fuzzy Query Evaluation, Fuzzy Sets and Systems 38 pp137-152

[8] U.Höhle, (1976), MaBe auf unscharfen Mengen, Zeitshrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 36 pp179-188

[9] P.Smets (1981), The Degree of Belief in a Fuzzy Event, Information Sciences 25 pp1-19

[10] L.A.Zadeh, (1968), Probability Measures of Fuzzy Events, Journal of Mathematical Analysis and Applications 23, 421-427

[11] L. A.Zadeh, (1978), Fuzzy Sets as a Basis for a Theory of Possibility, Fuzzy Sets and Systems 1, 3-28