Gradual properties vs. uncertainty : Fuzzy logic vs. possibilistic logic

Didier Dubois – Henri Prade

I.R.I.T. – C.N.R.S.
Université Paul Sabatier
118 route de Narbonne
31062 Toulouse Cedex – France

Philippe Smets

I.R.I.D.I.A.
Université Libre de Bruxelles
50 avenue F. Roosevelt - CP 194/6
1050 Bruxelles – Belgique

In a recent paper [4], as well as in the revised version printed in this issue [5], Ch. Elkan has questioned fuzzy logic and cast serious doubts on the reasons of its success, arguing that "fuzzy logic collapses mathematically to two-valued logic". In this note we want to formally express our complete disagreement with such a claim. We especially object to :

1) The use, in Ch. Elkan's proof, of too strong a notion of logical equivalence. The particular equivalence that he considers in his proof, and which is valid in Boolean algebra, has nothing to do with fuzzy logic ;
2) The claim that De Morgan's algebra "allows very little reasoning about collections of fuzzy assertions". However Ch. Elkan correctly states that when logical equivalence is restricted to De Morgan algebra equalities, "collapse to two truth values is avoided".

Furthermore, Ch. Elkan fails to understand the important distinction [9][11] between two totally different problems to which fuzzy set-based methods apply, namely the handling of *gradual* (thus non-Boolean) properties whose satisfaction is a matter of degree (even when information is complete) on the one hand, and the handling of uncertainty pervading Boolean propositions and induced by incomplete states of knowledge which are represented by means of fuzzy sets on the other hand. The first problem requires the plain use of fuzzy sets, while the second one is the realm of possibility theory [24][8] and possibilistic logic [10]. We now discuss in greater details these two above points and the distinction between truth functional fuzzy (multiple-valued) logic and non-fully compositional possibilistic logic.

Classical propositional logic equivalence does not hold in fuzzy logic

Ch. Elkan [4][5] claims that in fuzzy logic the four following requirements hold for any assertions A and B, t being a truth assignment function such that $\forall A$, $t(A) \in [0,1]$

$$t(A \wedge B) = \min(t(A), t(B)) \tag{1}$$

$$t(A \vee B) = \max(t(A),t(B)) \quad (2)$$
$$t(\neg A) = 1 - t(A) \quad (3)$$
$$t(A) = t(B) \text{ if A and B are logically equivalent.} \quad (4)$$

While (1)-(2)-(3) are indeed the basic relations governing degrees of truth in fuzzy logic (as well as fuzzy set membership degrees) as proposed by Zadeh [23], requirement (4) where "logically equivalent" is understood in a stronger sense than the equivalences induced by (1)-(2)-(3) has never been seriously considered by any author in the fuzzy set literature (up to a few erroneous papers which may always exist in a corpus of more than 10.000 published papers in the fuzzy set literature). Obviously some classical logic equivalences still hold with fuzzy assertions obeying (1)-(2)-(3), namely the ones allowed by the De Morgan's structure induced by (1)-(2)-(3), as for instance

$$A \wedge A \equiv A \ ; \ A \vee A \equiv A \ \text{(idempotency)}$$
$$A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C) \ ; \ A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C) \ \text{(distributivity).}$$

But other Boolean equivalences *do not* hold, for instance

$$A \wedge \neg A \# \perp \text{ since (1) and (3) } \textit{only} \text{ entail } t(A \wedge \neg A) = \min(t(A), 1 - t(A)) \leq 1/2$$
$$A \vee \neg A \# \top \text{ since (2) and (3) } \textit{only} \text{ entail } t(A \vee \neg A) = \max(t(A), 1 - t(A)) \geq 1/2$$

where $t(\perp) = 0$ and $t(\top) = 1$. Indeed the failure of contradiction and excluded-middle laws is typical of fuzzy logic as emphasized by many authors. This is natural with gradual properties like 'tall'. For instance, in a given context, somebody who is 1.75 meter high, may be considered neither as completely tall (i.e., tall with degree 1) nor as completely not tall (i.e., tall with degree 0) ; in this case we may have, for example, $\mu_{\text{tall}}(1.75) = 0.5 = \mu_{\neg\text{tall}}(1.75)$.

For establishing the collapse of fuzzy logic to binary logic, Ch. Elkan [4][5] uses the following logical equivalence

$$\neg(A \wedge \neg B) \equiv B \vee (\neg A \wedge \neg B) \quad (5)$$

postulated as being "plausible intuitively". The left hand part of (5) can be equivalently written in fuzzy logic, if (1)-(2)-(3) hold

$$\neg(A \wedge \neg B) \equiv \neg A \vee B$$

while the right hand part of (5) can be equivalently written in fuzzy logic, if (1)-(2)-(3) hold

$$B \vee (\neg A \wedge \neg B) \equiv (\neg A \vee B) \wedge (B \vee \neg B)$$

which clearly relates to the excluded-middle law ; thus it is expected that (5) fails to hold in fuzzy logic, and indeed it can be checked by applying (1)-(2)-(3) that a counter-example to (5) is provided by $t(A) = 0$, $t(B) = 0.5$ for instance. Thus Ch. Elkan's claim of "a paradox in fuzzy logic" relies only on faulty assumptions, or at best on a logical equivalence the rationale of which is far from natural in the scope of fuzzy logic.

Gradual and interpolative reasoning in fuzzy logic

As already said, fuzzy logic is concerned with the handling of assertions like 'John is tall', whose truth is a matter of degree due to the presence of gradual predicates in them. The degree of truth of compound expressions can be easily computed using (1)-(2)-(3). Note that although we restrict ourselves in this short discussion to the operators minimum, maximum, complement to one, there exists a panoply of other operators (see, e.g., [22][16]) which enables us to model different kinds of 'AND' and 'OR' operations between properties in a multi-criteria aggregation perspective.

More than tweety years ago, R.C.T. Lee [19] has provided the basic machinery for reasoning in fuzzy logic (in the sense of (1)-(2)-(3) by extending the resolution rule to fuzzy logic). He established that if all the truth values of the parent clauses are greater than 0.5, then a resolvent clause derived by the resolution principle always has a truth-value between the maximum and the minimum of those of the parent clauses.

We may also use an implication operator for modelling so-called gradual rules [14] which express pieces of knowledge of the form "the more X is A, the more Y is B", e.g., "the taller you are, the heavier you are". This is captured by the implication defined by

$$\begin{aligned} t(A \rightarrow B) &= 1 \text{ if } t(A) \leq t(B) \\ &= 0 \text{ if } t(A) > t(B). \end{aligned} \tag{6}$$

Note that this implication is the natural counterpart of Zadeh's fuzzy set inclusion defined by the pointwise inequality of the membership functions [23]. This implication is directly associated with (1)-(2)-(3) since $A \rightarrow B \equiv T$ iff $A \wedge B \equiv A$. Such an implication purely expresses a gradual relationship and has nothing to do with uncertainty.

Besides, Takagi and Sugeno [21] have proposed an interpolation mechanism between n rules with fuzzy condition parts and non-fuzzy conclusions of the form "if X is A_i and Y is B_i then Z = c_i", by computing the following output when X = x_0 and Y = y_0 is observed

$$Z = \frac{\sum_i \gamma_i \cdot c_i}{\sum_i \gamma_i} \tag{7}$$

where $\gamma_i = \min(\mu_{A_i}(x_0), \mu_{B_i}(y_0))$, i = 1,n. Again this kind of "inference" which is widely used in fuzzy control has nothing to do with uncertainty handling, since only an interpolation between typical conclusions is performed based on degrees of similarity between the input (x_0,y_0) and the prototypical values in the core of the fuzzy set $A_i \times B_i$. This similarity is measured by the coefficients γ_i which cannot be considered as degrees of uncertainty in any case. In spite of its apparently ad hoc nature, (7) can be justified using (6) and viewing the rules as expressing "the more X is A_i and Y is B_i, the *closer* Z is to c_i" and using appropriately shaped membership functions ; see [15] for details.

This shows that, contrarily to Ch. Elkan's claim, some worth-considering kinds of reasoning, as examplified by Lee and Takagi-Sugeno methods, can be handled in a De Morgan algebra framework.

Possibility theory and the handling of uncertainty

Fuzzy sets can be used not only for modelling the gradual nature of properties but can also be used for representing incomplete states of knowledge. In this second use, the fuzzy set plays the role of a possibility distribution which provides a complete ordering of mutually exclusive states of the world according to their respective levels of possibility or plausibility. For instance, if we *only know* that "John is tall" (but not his precise height), where the meaning of 'tall' is described in the context by the membership function of a fuzzy set, i.e., μ_{tall}, then the greater $\mu_{tall}(x)$ is, the greater the possibility that height(John) = x and the smaller $\mu_{tall}(x)$, the smaller this possibility.

Given a [0,1]-valued possibility distribution π describing an incomplete state of knowledge, Zadeh [24] defines a so-called possibility measure Π such that

$$\Pi(A) = \sup\{\pi(x), x \text{ makes A true}\} \tag{8}$$

where A is a *Boolean* proposition, i.e., a proposition which can be true or false only. It can be easily checked that for Boolean propositions A and B, we have

$$\Pi(A \vee B) = \max(\Pi(A),\Pi(B)) \tag{9}$$

but that we only have the inequality

$$\Pi(A \wedge B) \leq \min(\Pi(A),\Pi(B)) \tag{10}$$

in the general case (equality holds when A and B are *logically independent*). Indeed if $B \equiv \neg A$, $\Pi(A \wedge B) = \Pi(\bot) = 0$, while $\min(\Pi(A),\Pi(\neg A)) = 0$ only if the information is sufficiently complete for having either $\Pi(\neg A) = 0$ (A is true) or $\Pi(A) = 0$ (A is false). If nothing is known about A, we have $\Pi(A) = \Pi(\neg A) = 1$. By duality, a necessity measure N is associated to Π according to the relation (which can be viewed as a graded version of the relation between what is necessary and what is possible in modal logic)

$$N(A) = 1 - \Pi(\neg A) \tag{11}$$

which states that A is all the more necessarily true as $\neg A$ has a low possibility to be true. It entails

$$N(A \wedge B) = \min(N(A),N(B)) \tag{12}$$

$$\text{and} \quad N(A \vee B) \geq \max(N(A),N(B)). \tag{13}$$

The equalities (9), (11) and (12) should not be confused with (2), (3) and (1) respectively. In (9), (11), (12) we deal with Boolean propositions pervaded with uncertainty due to incomplete information, while (1)-(2)-(3) pertain to non-Boolean propositions whose truth is a matter of degree (the information being assumed to be complete). Very often, discussions about fuzzy expert systems or uncertain knowledge base systems get confused because of a lack of distinction between degrees of truth and degree of uncertainty. Fuzzy logic, as understood by Ch. Elkan, is a logic where the truth-status of propositions is multiple-valued ; i.e., intermediary truth values between true and false (like "very true", "rather true", etc.) exist. On the contrary degrees of uncertainty apply to all-or-nothing propositions, and do not model truth values but express the fact that the truth value (true or false) is unknown. The uncertainty degrees then try to assess which one of 'true' or of 'false' is the most plausible truth value. This distinction was already made by one of the founders of subjective probability theory, namely De Finetti [6], and has been quite forgotten since then in the AI community and by Ch. Elkan in particular. Still we consider that this distinction is a crucial prerequisite in any discussion about fuzzy sets and possibility theory and their use in automated reasoning.

Observe also that neither Π, nor N, are fully compositional with respect to $\wedge$, $\vee$ and $\neg$. This is not surprizing since the only way to map a Boolean structure on [0,1] by a fully compositional mapping f is to have f(A) equal to 0 or to 1 for any A, as pointed out in [9]. Truth-functionality in (1)-(2)-(3) is preserved only by having A and B elements of a weaker structure, namely a De Morgan algebra. Thus, logics of uncertainty cannot be fully compositional with respect to uncertainty degrees, a point that is also recognized by Ch. Elkan [5] in the case of probability measures, and which dates back at least to De Finetti [6] in the thirties ! However partial compositionality is possible ; probabilities are compositional with respect to negation, possibilities with respect to disjunction, necessities with respect to conjunction. However reading his paper, it seems that Ch. Elkan never heard about possibility theory, which is another side of fuzzy sets.

Let us consider Ch. Elkan's watermelon example as stated in [5]. We have

$$\text{watermelon}(x) \equiv \text{redinside}(x) \wedge \text{greenoutside}(x).$$

It is supposed that "for some melon m, evidence that m is red internally has strength 0.5 and m is green externally with strength of evidence 0.8". It is not clear what Ch. Elkan means by "strength of evidence" in the light of the above comments. We shall assume they are indeed degrees of uncertainty, rather than degrees of red and degrees of green. But then the only way to anchor this discussion in the fuzzy logic debate is to interpret these degrees in possibility theory. Ch. Elkan's watermelon sentence can be understood as N(redinside(m)) $\geq$ 0.5 and N(greenoutside(m)) $\geq$ 0.8, expressing that the available information makes us certain that m is red inside to the degree 0.5, and that it is green outside to the degree 0.8. A direct application of (12) leads to N(watermelon(m)) $\geq$ min(0.5, 0.8) = 0,5, a result also obtained under an equality form by Ch. Elkan by applying (1) in a unappropriate way. However Ch. Elkan [] would like to conclude that "m is a watermelon with strength of evidence over 0.5". This seems a strange requirement that a probabilistic model would not satisfy either (since Prob(A $\wedge$ B) $\leq$ min(Prob(A),Prob(B)). Indeed we are not here in a data fusion situation where two independent sources provide the same conclusion with various strengths (see [13] for an overview of data fusion issues using possibilistic and fuzzy models), but in a case where the logical conjunction of two conditions is required to conclude that m is a watermelon (namely the inside redness of m and its outside greenness). Note that in case we have both that N(A) $\geq \alpha$ and that N(A) $\geq \alpha'$ as obtained from distinct arguments, we shall conclude that N(A) $\geq$ $\max(\alpha,\alpha')$.

Reasoning under uncertainty with possibility theory

Possibilistic logic is a logic where first order logic formulas are weighted by lower bounds of necessity or possibility measures, which reflect the uncertainty of the available information. Possibilistic logic [10][11, Part 2] has been developed both at the syntactic level where an inference machinery based on extended resolution and refutation exists (the lower bound of the necessity of a resolvent clause is the minimum of the lower bounds of the necessity measures of the parent clauses), and at the semantic level where a semantics in terms of a possibility distribution over a set of classical interpretations has been proved to be sound and complete with respect to the syntax. Due to the fact that a possibility distribution encodes a preferential ordering over a set of possible interpretations, possibilistic logic has been shown to capture an important class of nonmonotonic reasoning consequence relations [12] and has capabilities for handling partial inconsistency in knowledge bases [10]. Moreover, possibilistic Assumption-based Truth Maintenance Systems [7] based on possibilistic logic have been defined for dealing with uncertain justifications and ranking environments in a label ; it was successfully applied to a data fusion application [20].

However, possibility theory offers more general applications to reasoning with uncertain, imprecise or fuzzy pieces of information by manipulating possibility distributions explicitly. An example of these reasoning capabilities is provided by the so-called generalized modus ponens [25] which from a fuzzy fact X is A' represented by a possibility distribution $\pi_X = \mu_{A'}$ and a fuzzy rule if X is A then Y is B also represented by a possibility distribution $\pi_{Y|X}$ enables us to infer the possibility distribution restricting the possible values of Y by combining π_X and $\pi_{Y|X}$ and projecting the result on the domain of the variable Y. According to the multiple-valued logic implication $\rightarrow$ used for computing $\pi_{Y|X}$ from μ_A and μ_B, different kinds of fuzzy rules can be modelled. In particular we can distinguish, for instance, between the purely gradual rules already mentioned (of the form "the more X is A, the more Y is B") and certainty rules of the form "the more X is A the more *certain* Y is B". Thus graduality can be also encountered in the expression of incomplete states of knowledge pertaining to ill-known relationships between variables (like the ones expressed by fuzzy rules). See [11] for more details.

Expert systems with fuzzy rules have been designed, which are not as simple as fuzzy controllers (where no chaining of rules is required but only an interpolation between the conclusions of a set of parallel rules). These expert systems, as expected by Ch. Elkan [5], do "knowledge-intensive tasks such as diagnosis, scheduling, or design". Here we will only mention a few representative references among others : CADIAG-2 [1], TAIGER [17], RUM [3], MILORD [18], OPAL [2]. All these systems have been or are still applied to genuine applications in one of the above-mentioned fields. These systems use some form of fuzzy set or possibility theory-based inference mechanisms which is much more sophisticated than the simple use of the three formulas (1)-(2)-(3) originally proposed by Zadeh in 1965 and to which fuzzy set and possibility theory methods cannot be reduced. See [11, Part 2] for a brief analysis of these systems. For the sake of brevity, we do not mention here many other noticeable works on fuzzy set and possibility theory-based inference systems which process temporal knowledge, belief networks or perform qualitative or abductive reasoning.

Concluding

The main message of this response to Ch. Elkan's papers [4][5] is that "fuzzy logic", using this popular expression here as a generic term, is not as simple as this author seems to believe (in that respect the absence of any mention in Ch. Elkan's discussion [4][5] of Zadeh's possibility theory and approximate reasoning approach [24][25] is quite a revealing sign).

In the literature the expression "fuzzy logic" usually refers either to multiple-valued logic (as is the first part of Elkan's paper), or to fuzzy controllers. However it should be clear that the two domains have very little in common, due to the fact that control engineers usually do not know about logic, and logicians do not know about control. In that sense the first part of Ch. Elkan's paper has very little relevance to his discussion on fuzzy control. If the success of fuzzy logic is paradoxical, it is certainly not because of Ch. Elkan's collapsing property. Besides Zadeh's view of fuzzy logic seems to go far beyond multiple-valued logic, and is as much a framework for handling incomplete information as a methodology for capturing graduality in propositions. The concept of fuzzy truth values refers as much to the idea of a partially unknown truth value as to intermediate truth values. This is why we have emphasized the crucial distinction between the truth-functional handling of gradual properties and the possibilistic treatment of uncertainty (which is not fully compositional).

It is certainly true that the huge amount of fuzzy set literature (whose quality may be unavoidably felt unequal) does not contribute to help newcomers to have a synthetic, well-informed and balanced view of the domain. See [] for a representative and topically organized sampling of papers published in the fuzzy set field in the last tweenty years. Similarly, the great success encountered by fuzzy controllers where a simple and efficient way of implementing an interpolative mechanism is provided, should not hide the other existing applications and more generally the great potential of fuzzy set and possibility theory for Artificial Intelligence applications.

References

[1] Adlassnig K.P., Kolarz G., CADIAG-2 : Computer-assisted medical diagnosis using fuzzy subsets. In : Approximate Reasoning in Decision Analysis (M.M. Gupta, E. Sanchez, eds.), North-Holland, Amsterdam, 1982, 219-247.

[2] Bensana E., Bel E., Dubois D., OPAL : A multi-knowledge-based system for industrial job-shop scheduling. Int. J. of Prod. Res., 26(5), 1988, 795-819.

[3] Bonissone P.P., Gans S.S., Decker K.S., RUM : A layered architecture for reasoning with uncertainty. Proc. of the 10th Inter. Joint Conf. on Artificial Intelligence (IJCAI'87), Milan, Italy, Aug. 23-28, 1987, 891-898.

[4] Ch. Elkan, The paradoxical success of fuzzy logic. Proc. of the 11th National Conf. on Artificial Intelligence (AAAI'93), Washington, D.C., July 11-15, 1993, 698-703.

[5] Ch. Elkan, The paradoxical success of fuzzy logic. IEEE Expert, this issue.

[6] De Finetti B., La logique de la probabilité. Actes du Congrès Inter. de Philosophie Scientifique, Paris, 1935, Hermann et Cie Editions, 1936, pp. IV1-IV9.

[7] Dubois D., Lang J., Prade H. A possibilistic assumption-based truth maintenance system with uncertain justifications, and its application to belief revision. In : Truth Maintenance Systems (Proc. of the ECAI'90 Workshop, Stockholm, Sweden, Aug. 1990) (J.P. Martins, M. Reinfrank, eds.), Lecture Notes in Artificial Intelligence, Vol. 515, Springer Verlag, Berlin, 1990, 87-106.

[8] Dubois D., Prade H. (with the collaboration of Farreny H., Martin-Clouaire R., Testemale C.), Possibility Theory — An Approach to Computerized Processing of Uncertainty. Plenum Press, New York, 1988.

[9] Dubois D., Prade H., An introduction to possibilistic and fuzzy logics. In : Non-Standard Logics for Automated Reasoning (Ph. Smets, E.H. Mamdani, D. Dubois, H. Prade, eds.), Academic Press, New York, 1988, 287-326.

[10] Dubois D., Prade H., Epistemic entrenchment and possibilistic logic. Artificial Intelligence, 50, 1991, 223-239.

[11] Dubois D., Prade H., Fuzzy sets in approximate reasoning. Fuzzy Sets and Systems, 40, 1991 : Part 1 : Inference with possibility distributions, 143-202 ; Part 2 (with Lang J.) : Logical approaches, 203-244.

[12] Dubois D., Prade H., Possibilistic logic, preferential models, non-monotonic and related issues. Proc. of the 12th Inter. Joint Conf. on Artificial Intelligence (IJCAI'91), Sydney, Australia, Aug. 24-30, 1991, 419-424.

[13] Dubois D., Prade H., Combination of fuzzy information in the framework of possibility theory. In : Data Fusion in Robotics and Machine Intelligence (M.A. Abidi, R.C. Gonzalez, eds.), Academic Press, New York, 1992, 481-505.

[14] Dubois D., Prade H., Gradual inference rules in approximate reasoning. Information Sciences, 61, 1992, 103-122.

[15] Dubois D., Prade H., Possibility theory as a basis for preference propagation in automated reasoning. Proc. of the 1st IEEE Inter. Conf. on Fuzzy Systems (FUZZ-IEEE'92), San Diego, CA, March 8-12, 1992, 821-832.

[16] Dubois D., Prade H., Yager R.R. (Eds.), Readings in Fuzzy Sets for Intelligent Systems. Morgan & Kaufmann, 1993.

[17] Farreny H., Prade H., Wyss E., Approximate reasoning in a rule-based expert system using possibility theory : A case study. In : Information Processing'86 (H.J. Kugler, ed.), North-Holland, Amsterdam, 1986, 407-413.

[18] Godo L., López de Mántaras R., Sierra C., Verdaguer A., MILORD : The architecture and the management of linguistically expressed uncertainty. Int. J. of Intelligent Systems, 4, 1989, 471-501.

[19] Lee R.C.T., Fuzzy logic and the resolution principle. J. Assoc. Comput. Mach., 19, 1972, 109-119.

[20] Monai F.F., Chehire T., Possibilistic assumption based truth maintenance system, validation in a data fusion application. Proc. of the 8th Conf. on Uncertainty in Artificial Intelligence (D. Dubois, M.P. Wellman, B. D'Ambrosio, Ph. Smets, eds.), Stanford, CA, July 17-19, 1992, Morgan & Kaufmann, 1992, 83-91.

[21] Takagi T., Sugeno M., Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. on Systems, Man and Cybernetics, 15, 1985, 116-132.

[22] Yager R.R., Connectives and quantifiers in fuzzy sets. Fuzzy Sets and Systems, 40, 1991, 39-75.

[23] Zadeh L.A., Fuzzy sets. Information and Control, 8, 1965, 338-353.

[24] Zadeh L.A., Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 1978, 3-28.

[25] Zadeh L.A., A theory of approximate reasoning. In : Machine Intelligence, Vol. 9 (J.E. Hayes, D. Mitchie, L.I. Mikulich, eds.), Wiley, New York, 1979, 149-194.