A Probabilistic Characterization of Fuzzy Set Membership, with Application to Mixed Fuzzy-Probabilistic Inference

Ben Goertzel
Novamente LLC

August 5, 2009

Abstract

A simple probabilistic grounding of the "fuzzy set membership degree" is presented, and used to provide definitions of the absolute and conditional probabilities of fuzzy sets. Among other possible applications, this allows fuzzy membership values to be coherently incorporated into probabilistic reasoning processes.

1 Introduction

Fuzzy set theory [3] and probability theory provide two qualitatively different, separately useful methods for quantifying uncertainty. In the Probabilistic Logic Networks reasoning framework [2], fuzzy and probabilistic truth values are interrelated within the same inference processes, and this interrelationship is argued to be essential to commonsense reasoning. However, [2] does not fully formalize this interrelationship; the purpose of this note is to remedy this gap, by giving a formulation of fuzzy truth values and fuzzy set theory in probabilistic terms.

It must be emphasized that the formulation given here does not constitute a reduction of fuzzy truth values to probabilities in a naive sense. For instance, we are not claiming that the fuzzy statement $\chi_F(s) = c$ is equivalent to the probabilistic statement $P(s \in F) = c$. Rather, our formulation attempts to capture the unique semantics of fuzzy set membership, via probabilities defined in terms of certain multisets derived from fuzzy sets.

One concrete result from this formulation is a firm probabilistic grounding for the PLN formulas given in [2] for the conditional probability $P(F|G)$ and the distance $d(F,G)$ between two fuzzy sets F and G. We also discuss the interoperation of fuzzy sets with probabilistic quantifiers, and sketch an approach to a general unification of fuzzy and probabilistic logics using fibring of logics.

2 A Probabilistic Characterization of Fuzzy Sets and their Interrelationships

Where N is a positive integer and S is a set of cardinality n, consider the space $\mathcal{F}$ of all fuzzy sets formed from the elements of S, whose membership degrees are restricted to the set $\{\frac{k}{N}, k = 0, ..., N\}$.

Construct a multiset $\mathcal{O}$ called the *observation set*, which contains $N\chi_F(s)$ copies of the pair (s, F) for each $F \in \mathcal{F}$ and $s \in S$.

Intuitively, to picture this multiset one may envision a "property detector" machine that, when used to scan an object s, prints out up to nN paper cards, each of which has some marking (s, F) on it. The number of cards marked (s, F) that are produced, may be interpreted as a scaled version of the degree to which s has property F (specifically, it's equal to this degree multiplied by N).

For $F \in \mathcal{F}$, define $P(F)$ as the probability that, if a random item is chosen from $\mathcal{O}$, it is a pair of the form (s, F) for some $s \in S$. I.e., it is the probability that a randomly chosen card, produced by the machine (when applied to a randomly chosen object), has (s, F) printed on it for some s.

It follows from this definition that

$$P(F) = \frac{\sum_{s \in S} \chi_F(s)}{\sum_{s \in S, H \in \mathcal{F}} \chi_H(s)}$$

Next, each $F \in \mathcal{F}$ may be identified with a multiset M_F of cardinality $\sum_{s \in S} \chi_F(s) \leq Nn$, which contains $N\chi_F(s)$ copies of each $s \in S$. The standard multiset intersection defines the number of copies of s in $F \cap G$ as the minimum of the number of copies of s in F and the number of copies of s

in G. Based on this definition, multiset intersection coincides with standard fuzzy set intersection, and we find

$$M_{F \cap G} = M_F \cap M_G$$

A distribution over these multisets may be imposed by setting $P(M_F) = P(F)$.

Thus if we have two fuzzy sets F and G drawn from $\mathcal{F}$, we can define

$$P(F|G) = \frac{P(F \cap G)}{P(G)}$$

or equivalently

$$P(F|G) = \frac{P(M_F \cap M_G)}{P(M_G)}$$

Relatedly, one can construct a similarity measure and a metric on the space of fuzzy sets via

$$sim(F, G) = \frac{P(F \cap G)}{P(F \cup G)}$$

$$d(F, G) = 1 - sim(F, G)$$

where the latter is easily seen to be equal to the Tanimoto distance [4] between the vectors defined from the membership functions of F and G.

3 A Few Extensions

3.1 The Weighted Case

A natural extension to the above is obtained by considering the case where the elements $s \in S$ have prior probabilities $\pi(s)$; then one may say

$$P(F) = \frac{\sum_{s \in S} \pi(s) \chi_F(s)}{\sum_{s \in S, H \in \mathcal{F}} \pi(s) \chi_H(s)}$$

3.2 Distributions over Fuzzy Sets

If instead of an individual fuzzy set F one has a distribution ν defined over $\mathcal{F}$, then one can define

$$P(F) = \frac{\sum_{s \in S, F} \nu(F) \chi_F(s)}{\sum_{s \in S, H \in \mathcal{F}} \chi_H(s)}$$

and proceed similarly to above. Of course one may also use a distribution π over S together with this distribution ν.

3.3 The Continuous Case

If S is infinite, and one looks at the set $\mathcal{F}$ of fuzzy sets with elements in S whose membership values live in the continuum between 0 and 1, then one can extend the above to assess the probability of a *set of fuzzy sets* $\mathbf{R} \subset \mathcal{F}$ by defining

$$P(\mathbf{R}) = \frac{\int_{s \in S, F \in \mathbf{R}} \chi_F(s) d\mu(s) d\nu(F)}{\int_{s \in S, H \in \mathcal{F}} \chi_H(s) d\mu(s) d\nu(H)}$$

for appropriate probability measures μ and ν; and under appropriate assumptions one finds that this continuous probability is a limit of the probabilities of discrete fuzzy sets as $N \to \infty$. Continuum-based conditional probabilities then follow in the obvious way; and one may add a distribution π over S if desired. However, we shall not focus on this case here.

4 More Complex Mixtures of Fuzzy and Probabilistic Operators

Beyond the simple case considered above, involving conditional probabilities among fuzzy sets, there are also more complex cases, in which one must mix fuzzy and probabilistic operators in moderately subtle ways to capture the semantics of commonsense relationships between fuzzy terms. The mathematics here becomes more involved, and in this section we will only sketch the main ideas that must be introduced to handle this case, leaving the details for another publication.

4.1 Mixing Fuzzy Functions and Probabilistic Quantifiers

Consider for instance the case

$$P(hasTallDad|tall)$$

Given in this form, this is formalized by the multiset approach described above. But what if we need to define $hasTallDad$ in terms of the primitives $tall$ and $dadOf$?

We can define

$$hasTallDad(x) = \exists y : ((x \in tall) \cap (dadOf(x, y)))$$

but the semantics of the right hand side requires some explanation.

Firstly, $\cap$ is a fuzzy conjunction, defined using the familiar min-formula from fuzzy set theory (which means that $dadOf(x, y)$ is interpreted as having a fuzzy truth value). Thus, $F(x, y) = ((x \in tall) \cap dadOf(x, y)))$ is a function that maps a pair (x, y) into a fuzzy truth value.

To interpret the existential quantifier, we set $F_x(y) = F(x, y)$, and note that if we assume a prior probability distribution over y values, then for fixed x this yields a probability distribution over $F_x(y)$ (fuzzy) values. Given this distribution, we can define an *indefinite truth value* for $\exists y : F_x(y)$ using the approach given in [2].

Finally, we consider $hasTallDad(x)$ as a fuzzy set where $\chi_{hasTallDad}(x)$ is determined by the indefinite truth value of $\exists y : F_x(y)$. An indefinite truth value is a confidence interval $([L, U], b)$ (where b is the confidence level, L is the lower and U the upper bound, and there is also a parameter k), so if we want $hasTallDad(x)$ to have an ordinary single-number fuzzy truth value, we need to project the indefinite truth value into a single number, such as the midpoint of L and U.

4.2 Fibring Fuzzy and Probabilistic Logic

To formalize generally what this example illustrates, one route is to fibre together probabilistic and standard fuzzy logic [1]. Setting quantitative truth value formulas aside, one can construct a homogeneous fibring of the mixed predicate/term logic from [2] and fuzzy logic (considering the two logics to contain entirely different operator sets, e.g. considering fuzzy conjunction to be a different operation from probabilistic conjunction): syntactically, this is simply the logic whose set of inference rules is the union of the set of PLN inference rules and the set of fuzzy inference rules. Then, one can derive uncertain truth value formulas for this fibred logic by:

- applying probabilistic truth value formulas to PLN inference rules and fuzzy truth value formulas to fuzzy inference rules
- using the approach from Section 2 to deal with (PLN) logical implications between fuzzy terms
- using the method from the $hasTallDad$ example above to deal with (PLN) quantifiers over fuzzy terms

While a full formalization of this approach, would go beyond the scope of this paper (as it would require consideration of the axiomatization of PLN), we hope the idea is clear from these brief comments; and in fact the detailed elaboration is fairly straightforward.

Acknowledgments Thanks are due to Nil Geisweiller, Matt Ikle' and Yan King Yin for valuable discussions that helped greatly in refining the ideas in this paper.

References

[1] Caleiro, C., A. Sernadas and C. Sernadas (2005). Fibring logics: past, present and future. In Artemov et al. (ed.) We Will Show Them: Essays in Honour of Dov Gabbay, College Publications.

[2] Goertzel, Ben, Matthew Ikle, Izabela Freire Goertzel and Ari Heljakka (2008). Probabilistic Logic Networks. Springer Verlag.

[3] Klir, George and Bo Yuan (1995). Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall.

[4] Lipkus, A.H. (1999). A proof of the triangle inequality for the Tanimoto distance. Journal of Mathematical Chemistry, Volume 26, Numbers 1-3, 1999 , pp. 263-265(3)

[5] Syropoulos, Apostolos (2001) "Mathematics of Multisets" in C. S. Calude et al., eds., Multiset processing: Mathematical, computer science, and molecular computing points of view, LNCS 2235. Springer-Verlag: 347-358.