PROBABILITY LOGIC: SEMANTICS

by
Giangiacomo Gerla
Dipartimento di Matematica e Fisica dell'Università di Camerino, Italy.

1 - Introduction.

I will propose an approach to probability logic in which fuzzy logic is the basic tool. So, a suitable set of fuzzy rules of inferences is defined in such a way that the related fuzzy theories have a probabilistic meaning. Namely, by assuming that the set of formulas is a Boolean algebra **B**, a consistent probabilistic theory turns out to be a lower envelope (that is a meet of a family of probabilities) while the complete probabilistic theories coincide with the finitely additive probabilities. A promising possibility of probability logic is related to the management of probabilistic information in expert system theory.

While the semantic part of probability logic is exposed in this paper, the syntactical part will be given in [3].

2 - Fuzzy logic

Recall that, given a set F and a complete lattice L, an *L-subset* of F is any element of the direct power L^F, i.e. any map $s:F\rightarrow L$. The class of the L-subsets of F inherits the structure of a complete lattice by L and it extends the lattice of the subsets of F. Indeed if we call *crisp* an L-subset such that $s(x)\in\{0,1\}$ for every $x\in F$, then we can identify the subsets of F with the crisp L-subsets of F via the characteristic functions. We extend to the lattice L^F the terminology of set theory; for example if $s' \leq s$, we say the *s' is enclosed in s* or that *s' is a part of s* ; we call *union* and *intersection* the join and the meet operators, and so on. We say that s is *finite* if its support $Supp(s)=\{x \in F / s(x) \neq 0\}$ is finite. If L is the real interval [0,1], then we call *fuzzy subsets* the L-subsets; moreover, the *complement* $\sim s$ of a fuzzy subset s is defined by $(\sim s)(x)=1-s(x)$ for every $x \in F$.

The fuzzy logics are defined as follows. Let F be a set whose elements are called *formulas* , then a *semantics* in a fuzzy logic is any class S of L-subsets of F. We say that an element m of S is a *model* of an L-subset of formulas v and we write $m \models v$ provided that $v \leq m$. We call *satisfiable* any L-subset of formulas admitting a model in S. As an example, in multivalued sentential calculus S is the class of the valuations of the formulas; in classical first order logic, we can set S equal to the (characteristic functions) of the complete theories.

A semantics S induces a *logical consequence* operator $C_S : L^F \to L^F$ defined by setting, for every L-subset v of formulas

$$C_S(v)=\bigwedge\{m \in S / m \models v\}.$$

In particular if v is not satisfiable then $C_S(v)$ collapses in the whole set F of formulas, i.e. the map constantly equal to 1.

The syntactical apparatus is defined as follows: a *fuzzy rule of inference* is a pair $r=(r',r'')$ where r' is an n-ary operation defined in a subset $D_{r'}$ of F^n and r'' is an n-ary operation on L preserving joins in each variable. A *fuzzy syntax* on F is a pair (a,R) where a is an L-subset of F, the *L-subset of logical axioms* , and R is a set of fuzzy rules of inference. An L-subset s of formulas is *closed with respect to* the rule r if,

$$s(r'(\alpha_1,\dots,\alpha_n)) \geq r''(s(\alpha_1),\dots,s(\alpha_n)), \quad \text{for every} \quad (\alpha_1,\dots,\alpha_n) \in D_{r'}.$$

A *theory* on the fuzzy sintax (a,R) is an L-subset of formulas containing the L-subset of logical axioms and closed with respect to every rule in R. Obviously, the (characteristic function of the) whole set of formulas F is a theory that we call the *inconsistent theory*. We say that a theory is *complete* if it is a maximal element in the class of the consistent theories. A proof *of a formula* α is a sequence $\pi=\alpha_1,\dots,\alpha_m$ of formulas such that $\alpha_m=\alpha$ equipped with related "justifications". That is, for every $i=1,\dots,m$ we have to specify wheether i) α_i is assumed as a logical axiom or; ii) α_i is assumed as a proper axiom or; iii) α_i is obtained by a rule. In the last case we have to indicate also the rule and the formulas in $\alpha_1,\dots,\alpha_{i-1}$ used to obtain α_i. A precise definition is the following. As usual, for every set X we denote by X^+ the free semigroup generated by X. Let $\Gamma=F\cup(F\times\{0\})\cup(F\times R\times N^+)$ and set , for every $w\in\Gamma$, $\ulcorner w=w$ if $w\in F$ and $\ulcorner w=x$ if $w\in(F\times\{0\})$ or $w\in(F\times R\times N^+)$ and x is the first coordinate of w. An *R-proof of a formula* α is an element $w=w_1\dots w_m$ of Γ^+ such that $\alpha=\ulcorner w_m$ and, if $w_k=(x,r,i_1,\dots,i_n)$ then the rule r is n-ary, $k>1$, $1\le i_1,\dots,i_n\le k$ and $x=r'(\ulcorner w_{i_1},\dots,\ulcorner w_{i_n})$. Obviously, if $w=w_1\dots w_m$ is an R-proof, then $w(k)=w_1\dots w_k$, with $k\le m$, is an R-proof also.

Given an L-subset v of formulas, that we may interpret as a set of proper axioms, the *valuation* $Val(w,v)$ *of the R-proof w with respect to v* is completely defined by the equations

$Val(w,v) = v(w_m)$ if $w_m\in F$; $Val(w,v)=a(\ulcorner w_m)$ if $w_m\in F\times\{0\}$;

$Val(w,v)=r''(Val(w(i_1),v),\dots,Val(w(i_n),v))$ if $w_m=(x,r,i_1,\dots,i_n)$.

In [8] J.Pavelka proves the following facts.

Proposition 2.1 The intersection of a family of theories is a theory. Consequently, given an L-subset v of formulas, it is possible to define the *theory C(v) with L-subset of axioms v* by

(2.1) $$C(v) = \bigwedge\{\tau \ / \ \tau \text{ is a theory and } \tau\ge v\}$$

Moreover, we have

(2.2) $C(v)(\alpha) = \bigvee\{Val(w,v) \ / \ w \text{ is a proof of } \alpha\}$.

We call *deduction operator* the operator $C:L^F \to L^F$ defined by (2.1) and we say also that C(v) is *the L-subset of consequences of v.* We say that v is *inconsistent* if C(v) is the inconsistent theory.

A fuzzy logic is defined by a fuzzy semantics S and a fuzzy syntax (a,R). Obviously, it should be desirable to have a "completeness theorem", that is the coincidence between the logical consequence operator and the deduction operator.

3 - Probability logic: the semantics

In the sequel we assume that the set of formulas is a Boolean algebra **B** whose minimum and maximum we denote by **0** and 1, respectively. As an example, B could be the Lindenbaum algebra of a logic, an algebra of events and so on. The valuation lattices coincide with [0,1] and, as is usual in probability logic, the models are the probabilities defined in B, i.e. the finitely additive maps $p:\mathbf{B}\to[0,1]$ such that p(1)=1. Consequently, if IP denotes the set of probabilities in **B**, the semantic operator C_{IP} is defined by

$$C_{IP}(v) = \bigwedge\{p / p\in IP, \ p\geq v\}.$$

The map $C_{IP}(v)$ is not a probability, in general, so we are induced to consider the class of the fuzzy subsets of formulas that are intersections of probabilities. These maps are known in literature under the name of (lower) *envelopes* .

It is possible to give the following meaning to the fuzzy subset v of proper axioms . Assume that the (incomplete) informations we have about a random phenomenon enable us to find, for every formula α, only a lower bound $l(\alpha)$ and an upper bound $u(\alpha)$ to the unknown probability $p(\alpha)$, that is an interval approximation $[l(\alpha),u(\alpha)]$ of $p(\alpha)$.

If we define $v:B\to[0,1]$ by setting $v(\alpha)=l(\alpha)\vee(1-u(-\alpha))$, then $p(\alpha)\in[l(\alpha),u(\alpha)]$ for every formula α if and only if $p(\alpha)\geq v(\alpha)$. This means that it is possible to express our informations by saying that the unknown probability distribution p is a model of the fuzzy set of axioms v. In this case the meaning of $v(\alpha)$ is "the unknown probability of α is at least $v(\alpha)$". Conversely, every map $v:B\to[0,1]$ we interpret as above gives an interval approximation $[v(\alpha),1-v(-\alpha)]$ of $p(\alpha)$. Thus to give a fuzzy set of axioms in probability logic is equivalent to give an interval approximation of the unknown probability distribution.

Obviously, it is also possible that v is not satisfiable, that is no probability p exists such that $p(\alpha)\in[v(\alpha),1-v(-\alpha)]$. The next proposition shows the model compactness of probability logic.

Proposition 3.1. A fuzzy subset of formulas $v:B\to[0,1]$ is satisfiable if and only if every finite part of v is satisfiable .

For every h and k in N, we call *h-k-connective* the h-ary operation C^k defined by

$$C^k(\alpha_1,\dots,\alpha_h)=\bigvee\{\alpha_{i_1}\wedge\dots\wedge\alpha_{i_k} \,/\, i_1,\dots,i_k \text{ are distinct}\}$$

We have

$C^1(\alpha_1,\dots,\alpha_h)=\alpha_1\vee\dots\vee\alpha_h$, $C^2(\alpha_1,\dots,\alpha_h)=\bigvee\{\alpha_i\wedge\alpha_j / i\neq j,\ i,j\in\{1,\dots,h\}\}$,

$C^h(\alpha_1,\dots,\alpha_h)=\alpha_1\wedge\dots\wedge\alpha_h$, $C^{h+1}(\alpha_1,\dots,\alpha_h)=C^{h+2}(\alpha_1,\dots,\alpha_h)=\dots=0$;

The connection between the h-k-connectives and the probabilities is expressed by the following obvious proposition.

Proposition 3.2. A map $p:B\to[0,1]$ is a probability if and only if

(3.1) $$p(\alpha_1)+\dots+p(\alpha_h)=p(C^1(\alpha_1,\dots,\alpha_n))+\dots+p(C^h(\alpha_1,\dots,\alpha_h))$$

for every $\alpha_1,\dots,\alpha_h$ in B.

Also, we set

$$d(\alpha_1,\ldots,\alpha_h)=\min\{k\in N \;/\; C^k(\alpha_1,\ldots,\alpha_h)=0\}-1.$$

Obviously, $d(\alpha_1,\ldots,\alpha_h)=0$ if and only if $\alpha_1,\ldots,\alpha_h=0$ while otherwise

$$d(\alpha_1,\ldots,\alpha_h)=\max\{k\in N \;/\; C^k(\alpha_1,\ldots,\alpha_h)\neq 0\} .$$

Proposition 3.3 A fuzzy subset v of formulas is satisfiable if and only if, for every $\alpha_1, \ldots ,\alpha_h$ in **B**

(3.2) $$v(\alpha_1)+ \ldots +v(\alpha_h) \leq d(\alpha_1,\ldots,\alpha_h).$$

In Proposition 3.3 we can confine ourselves to formulas in Supp(v), then, if Supp(v) is finite, we can test easily condition (3.2).

Notice that from Proposition 3.3 it is possible to derive in a simple way a famous result in probability theory. Namely, every probability $p^*:B^*\to[0,1]$ defined in a subalgebra **B*** of **B** can be extended to a finitely additive probability defined in the whole algebra **B**. Indeed, let v be the extension of p* to **B** obtained by setting $v(x)=0$ for every $x\in B-B^*$. Then, since v satisfies (3.2) in its support, there exists a probability p defined in **B** such that $p\geq v$. Since $p^*(x)=v(x)=p(x)$ for every $x\in B^*$, p is an extension of p* .

Remark . The following stake interpretation of Proposition 3.3 is possible. Consider a game such that

- a player can bet on the occurrences of the events in **B** ;
- for every $\alpha\in B$, betting on α costs $v(\alpha)\in[0,1]$;
- the bank pay a fix stake of one unity .

It is also possible to bet on several different events. Obviously, if a player will bet on the (not necessarily distinct) events $\alpha_1,\ldots,\alpha_h$ then he has to pay $v(\alpha_1)+ \ldots +v(\alpha_h)$. Now, given a betting function v, we call *reasonable* any bet $\alpha_1,\ldots,\alpha_h$ in which there is the possibility of receiving at least the amount of money that the player has disbursed.

Thus, since it is easy to see that $d(\alpha_1,\ldots,\alpha_h)$ represents the maximum deal of money a player betting on $\alpha_1,\ldots,\alpha_h$ can win, Proposition 3.3 says that *v is satisfiable if and only if every bet in the related game is reasonable* . Obviously, the term "reasonable" refers to the player's point of view and not to the bank's point of view. For example, if v is constantly equal to zero every bet is reasonable for the player but no bet is reasonable for the bank.

As an example, due to the zero rule, the betting function of the roulette game is unsatisfiable.

The following proposition gives some characterizations of the envelopes. At first we define the analogous of the extension of a theory by a formula. Namely, given a fuzzy subset p of formulas and $\alpha \in B$, we denote by p_α, the *extension of p via* α , the fuzzy subset defined by $p_\alpha(x)=p(x)$ if $x \neq \alpha$ and $p_\alpha(x)=\max\{p(\alpha),1-p(-\alpha)\}$ if $x=\alpha$.

Proposition 3.4 Let p be a fuzzy subset of formulas such that $p(1)=1$, then the following are equivalent

a) p is an envelope;

b) $p(1)=1$ and $\sum p(\alpha_i) \leq (d-k+1)p(C^k(\alpha_1, \ldots, \alpha_h))+k-1$ for every $\alpha_1, \ldots, \alpha_h \in B$, $k \leq d$ and $d=d(\alpha_1,\ldots,\alpha_h)$;

c) every extension p_α of p via a formula α is satisfiable;

d) for every formula β there is a probability $p^* \geq p$ such that $p^*(\beta)=p(\beta)$.

REFERENCES

[1] BIACINO L. (1990), Generated envelopes, unpublished paper.

[2] BIACINO L., GERLA G. (1989), Decidability, recursive enumerability and Kleene hierarchy for L-subsets, *Zeitschr f. math. Logik und Grundlagen d. Math.* 35, 49-62.

[3] GERLA G. , Probability logic:syntax, to appear on this Journal.

[4] SCOTT D., KRAUSS P. (1966), Assigning probabilities to logical formulas, in "Aspects of Inductive Logic", North-Holland.

[5] DEMPSTER A.P. (1967), Upper and lower probabilities induced by a multivalued mapping, *Ann.Math.Statist.* 38, 325-339.

[6] DUBOIS D., PRADE H.(1988), Possibility Theory, Plenum Press, New York and London.

[7] PAPAMARCOU A., FINE T.L. (1986), A note on undominated lower probabilities, *The Annals of Probability* , 14, 710-723.

[8] PAVELKA J. (1979), On fuzzy logic I: Many-valued rules of inference, *Zeitschr.f.math. Logik und Grundlagen d. Math.* , 25, 119-134.

[9] SHAFER G. (1976), A Mathematical Theory of Evidence, Princeton University Press, Princeton.

[10] SUPPES P., ZANOTTI M. (1989), Conditions on upper and lower probabilities to imply probabilities, *Erkenntnis* 31, 323-345.