ON CONDITIONAL POSSIBILITY DISTRIBUTIONS

by

H. T. Nguyen

Memorandum No. UCB/ERL M77/52

11 August 1977

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ON CONDITIONAL POSSIBILITY DISTRIBUTIONS*

Hung T. Nguyen†

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California
Berkeley, California 94720

Abstract

In a recent paper in this Journal, L.A. Zadeh has defined the concept of a conditional possibility distribution. In the present paper, we show that, in order to be consistent with the notion of noninteraction of fuzzy variables, the expression for conditional possibility distribution must be normalized. A comparison of the properties of conditional possibility and probability distributions is made, and an application to the optimization of a possibilistic finite-state system is outlined.

---

*Research sponsored in part by the National Science Foundation Grant MCS76-07568 and U.S. Army Research Office Grant DAHC04-75-G0056.

†Current affiliation: Department of Mathematics and Statistics, University of Massachusetts, Amherst, Mass. 01002.

1. Introduction

In a recent paper [1], L.A. Zadeh has introduced the concept of a possibility distribution and presented a preliminary account of some of its basic properties.

Formally, the possibility distribution associated with a fuzzy variable is analogous to the probability distribution of a random variable and, like the latter, gives rise to the concepts of conditional and marginal possibility distributions, as described in [1]. The main result of our note is that, in order to achieve consistency with the concept of noninteraction of fuzzy variables, it is necessary to normalize the conditional possibility distribution in the manner described in Section 3. In addition, in Section 4 we outline an application of the concept of a conditional possibility distribution to the control of a possibilistic finite-state system -- an application which is suggestive of other possible applications of this concept to problems in which the variables are associated with possibility rather than probability distributions.

2. Possibility Distributions as Set Functions

Let X be a variable taking values in a universe of discourse U. As defined in [1], a possibility distribution, Π_X, associated with X is a fuzzy relation which acts as an elastic constraint on the values that may be assigned to X. Thus, if u is an element of U, then, by definition,

$$\pi_X(u) \triangleq \text{Poss}\{X = u\} \tag{1}$$

where $\pi_X(u)$, the possibility distribution function, is the membership function of Π_X and $\text{Poss}\{X = u\}$ is the possibility that X may take the value u, with the understanding that the function π_X is defined subjectively. A discussion of the connection between Π_X and the information conveyed by a fuzzy proposition may be found in [1].

The definition of π_X suggests that the possibility measure of a subset A of U be defined as an extension of (1) to subsets of U, i.e.,[1]

$$\pi_X(A) = \sup_{u \in A} \pi_X(u) \tag{2}$$

An immediate consequence of (2) is that $\pi_X(U) = 1$ if Π_X is normal and $\pi_X(U) < 1$ if Π_X is subnormal. Another direct consequence of (2) may be stated as the proposition:

<u>Proposition</u>. Given $\pi\colon U \to [0,1]$ such that

$$\sup_u \pi(u) = 1 \qquad (\text{i.e., } \Pi \text{ is normal})$$

then, for any possibility measure $\hat{\pi}$ (which is a mapping from the power set of U to [0,1]) such that

$$\hat{\pi}(\emptyset) = 0\ , \qquad \hat{\pi}(U) = 1$$

we have, for any index set I,

$$\pi(\bigcup_I A_i) = \sup_I \hat{\pi}(A_i) \tag{3}$$

For our purposes, we note that:

(a) $\hat{\pi}$ is an increasing function, i.e., $A \subset B \Rightarrow \hat{\pi}(A) \leq \hat{\pi}(B)$.

(b) The subadditivity expressed by (3) is analogous to the relation between a metric and an ultrametric.

(c) $\hat{\pi}(\{u\}) = \pi(u)$ is not, in general, identically zero.

---

[1] The concept of a possibility measure may be viewed as a special case of the more general concept of a fuzzy measure defined by Sugeno and Terano [10,11].

(d) If U is a Hausdorff topological space and π is an upper semi-continuous function, then $\hat{\pi}$ is a Choquet capacity which is, formally, a function of the inverse of some generalized information measure in the sense of Kampe de Feriet [5,6].

Now, if A is a fuzzy rather than nonfuzzy subset of U, what is the meaning that could be assigned to Poss{X is A}? In [1], it is proposed that Poss{X is A} be defined as

$$\text{Poss}\{X \text{ is } A\} = \sup_U [\mu_A(u) \wedge \pi_X(u)] \tag{4}$$

where μ_A is the membership function of A. We shall proceed to justify this definition. First, note that if A is nonfuzzy, then by (1)

$$\text{Poss}\{X \in A\} = \hat{\pi}(A) = \sup_A \pi_X(u) \ . \tag{5}$$

When A is fuzzy, (5) may be viewed as requiring the maximization of the real-valued function $\pi_X(\cdot)$ over the fuzzy constraint A [2]. In this connection, let $b(U,\mathbb{R})$ be the space of real-valued and bounded functions defined on U. If $f \in b(U,\mathbb{R})$, we write $S(f)$, $I(f)$ for $\sup_U f(u)$ and $\inf_U f(u)$, respectively.

<u>Definition 1</u>. The maximizing set of f [7] is the fuzzy subset $M(f)$ of U characterized by

$$\mu_{M(f)}(u) = \frac{f(u) - \beta(f)}{\alpha(f)} \tag{6}$$

where

$$\alpha(f) = S(f) \vee 0 - I(f) \wedge 0 \tag{7}$$

$$\beta(f) = S(f) \wedge 0 + I(f) \wedge 0 \tag{8}$$

($\vee$ and $\wedge$ stand for max and min respectively).

The minimizing set of f is the fuzzy subset $m(f)$ of U characterized by

$$\mu_{m(f)}(u) = \mu_{M(-f)}(u) \ . \tag{9}$$

<u>Definition 2</u>. For $f \in b(U,\mathbb{R})$ and $A \in \underset{\sim}{P}(U)$ (set of fuzzy subsets of U), the restriction of f to A, associated with the maximizing set $M(f)$, is defined as

$$f_A(u) = \alpha(f)\mu_{AM(f)}(u) + \beta(f) \ , \qquad u \in S_A \tag{10}$$

where $AM(f) = A \cap M(f)$, i.e.

$$\mu_{AM(f)}(u) = \mu_A(u) \wedge \mu_{M(f)}(u) \tag{11}$$

and $S_A = \{u | \mu_A(u) \neq 0\}$. The function $f_A(\cdot)$ can be written as

$$f_A(u) = f(u) \wedge \phi_{(f,A)}(u) \tag{12}$$

where $\phi_{(f,A)}(u) = \alpha(f)\mu_A(u) + \beta(f)$.

<u>Remark</u>. In the case of the minimizing set, $m(f)$, we define

$$\hat{f}_A(u) = -\alpha(f)\mu_{Am(f)}(u) + \beta'(f) \ , \qquad u \in S_A \tag{13}$$

where $\beta'(f) = S(f) \vee 0 + I(f) \vee 0$.

<u>Definition 3</u>. By the supremum of f over A, we mean the expression

$$S(f,A) = \sup_{S_A} f_A(u) \ . \tag{14}$$

In the same way, the infimum of f over A is expressed as

$$I(f,A) = \inf_{S_A} \hat{f}_A(u) \ . \tag{15}$$

It is shown in [2] that $S(f,A)$ and $I(f,A)$ have all of the basic properties of ordinary supremum and infimum (i.e., over nonfuzzy sets).

Remark. The motivation for defining $S(f,A)$ and $I(f,A)$ as above is to provide a general formulation for the optimization of real-valued functions under elastic constraints.

Now, if $f = \pi_X: U \to [0,1]$, then it is easy to check that

$$\mathrm{Poss}\{X \text{ is } A\} = \sup_{u \in S_A} f_A(u) = \sup[\mu_A(u) \wedge f(u)] \tag{16}$$

since $f_A(u) = \mu_A(u) \wedge f(u)$ in this case.

3. Conditional Possibility Distributions

General Considerations

Let $T: [0,1] \times [0,1] \to [0,1]$ be given, and let (X,Y) be a variable taking values in $U \times V$. Suppose that we can associate with (X,Y) some function $f_{(X,Y)}$ defined on $U \times V$ and taking values in the unit interval $[0,1]$. Then, in terms of this function, we define (or infer) the marginal distributions

$$f_X(u) = \theta_v[f_{(X,Y)}(u,v)] , \quad f_Y(u) = \theta_u[f_{(X,Y)}(u,v)] \tag{17}$$

where θ_v denotes some specific operation on v.

Based on T, we say that X and Y are T-independent iff

$$f_{(X,Y)}(u,v) = T[f_X(u),f_Y(v)] , \quad \forall (u,v) \in U \times V . \tag{18}$$

Note that we have to have the following consistency condition

$$\theta_v[T(f_X(u),f_Y(v))] = f_X(u) , \quad \forall u \in U . \tag{19}$$

Next, the conditional distribution is expressed as

$$f_{X|Y}(u|v) = f_{(X,Y)}(u,v)\cdot\alpha[f_X(u),f_Y(v)] \tag{20}$$

where the normalization function $\alpha(\cdot,\cdot)$ is a mapping from $[0,1]\times(0,1]\to\mathbb{R}^+$ such that $f_{X|Y}\leq 1$ and the following consistency condition holds

$$f_{X,Y}(u,v) = T[f_X(u),f_Y(v)] \Rightarrow f_{X|Y}(u|v) = f_X(u)\ , \quad \forall u\in U\ . \tag{21}$$

<u>Remark</u>. The essential idea here is that $\alpha(\cdot,\cdot)$ is a function of $f_Y(v)$ and $f_X(u)$ and not only of $f_Y(v)$, as is suggested by analogy with probability theory.

Now, the notion of non-interaction of two elastic constraints A and B implies that

$$\pi(A\cap B) = \pi(A)\wedge\pi(B) \tag{22}$$

which corresponds to $T(x,y) =$

$$T(x,y) = x\wedge y\ , \quad x,y\in[0,1]\ .$$

a) <u>Marginal possibility distributions</u> [1]

Taking

$$\theta_v(\cdot) = \sup_{v\in U}(\cdot) \tag{23}$$

leads to

$$\pi_X(u) = \sup_V \pi_{(X,Y)}(u,v)$$

from which it follows at once that for the case where π_X and π_Y are normal

$$\pi_X(u) = \sup_V[\pi_X(u)\wedge\pi_Y(v)]\ . \tag{24}$$

b) Conditional possibility distribution

The function $\alpha: [0,1] \times (0,1] \to \mathbb{R}^+$ must be such that

(i) $\pi_{(X,Y)}(u,v) \cdot \alpha[\pi_X(u), \pi_Y(v)] \in [0,1], \quad \forall u,v.$ (25)

(ii) $[\pi_X(u) \wedge \pi_Y(v)] \cdot \alpha[\pi_X(u), \pi_Y(v)] = \pi_X(u), \quad \forall u \in U.$ (26)

Solution. Consider the functional equation

$$(a \wedge b) \cdot \alpha(a,b) = a \tag{27}$$

for $a \in [0,1]$, $b \in (0,1]$.

The unique continuous solution (in a) of this equation is given by

$$\alpha(a,b) = \begin{cases} \dfrac{a}{a \wedge b} & \text{for } a \neq 0 \\ 1 & \text{for } a = 0 \end{cases}$$

or, in a more compact form,

$$\alpha(a,b) = \frac{a \vee b}{b} . \tag{28}$$

It will be shown below that this is the unique choice of $\alpha(\cdot,\cdot)$ such that the proposed expression for the conditional possibility distribution is consistent with the definition of non-interaction. First, we note that from

$$\alpha(a,b) = \begin{cases} 1 & \text{if } a \leq b \\ \dfrac{a}{b} & \text{if } a > b \end{cases}$$

it follows that $\alpha(a,b) \geq 1$.

Now

$$\pi(u|v) = \pi_{(X,Y)}(u,v) \cdot \alpha[\pi_X(u), \pi_Y(v)] \leq 1 \tag{29}$$

and

$$\pi_X(u) \le \pi_Y(v) \;\Rightarrow\; \pi(u|v) = \pi_{(X,Y)}(u,v) \le \pi_X(u) \wedge \pi_Y(v) \le 1 \tag{30}$$

$$\pi_X(u) > \pi_Y(v) \;\Rightarrow\; \pi(u|v) = \pi_{(X,Y)}(u,v) \cdot \frac{\pi_X(u)}{\pi_Y(v)} \tag{31}$$

$$= \frac{\pi_{(X,Y)}(u,v)}{\pi_Y(v)} \cdot \pi_X(u) \le \pi_X(u) \le 1$$

since $\pi_Y(v) = \sup_U \pi_{(X,Y)}(u,v)$.

It is easy to see that, if

$$\pi_{(X,Y)}(u,v) = \pi_X(u) \wedge \pi_Y(v)$$

then

$$\pi(u|v) = \begin{cases} \pi_X(u) \wedge \pi_Y(v) = \pi_X(u) & \text{if } \pi_X(u) \le \pi_Y(v) \\ \pi_Y(v) \cdot \dfrac{\pi_X(u)}{\pi_Y(v)} = \pi_X(u) & \text{if } \pi_X(u) > \pi_Y(v) \end{cases}$$

Thus, we are led to the definition of the conditional possibility distribution expressed by

$$\pi(u|v) = \begin{cases} \pi_{(X,Y)}(u,v) & \text{if } \pi_X(u) \le \pi_Y(v) \\ \pi_{(X,Y)}(u,v) \cdot \dfrac{\pi_X(u)}{\pi_Y(v)} & \text{if } \pi_X(u) > \pi_Y(v) \end{cases} \tag{32}$$

Remark. It should be observed that the value of $\pi(u|v)$ depends not only on $\pi_{(X,Y)}(u,v)$ and $\pi_Y(v)$, but also on $\pi_X(u)$ or, more specifically, on $\pi_{(X,Y)}(u,v)$ and the ratio of $\pi_X(u)$ to $\pi_Y(u)$.

We now proceed to derive a relation between $\pi_X(u)$, $\pi_Y(v)$ and $\pi(u,v)$, namely,

$$\pi_X(u) = \bigvee_{v \in V} [\pi(u|v) \wedge \pi_Y(v)] \; , \tag{33}$$

which is analogous to a corresponding relation for probabilities in probability theory.

First, we note that

$$\pi_{(X,Y)}(u,v) \le \pi(u|v) \le \pi_X(u) \;\Rightarrow\; \pi_X(u) = \sup_V \pi_{(X,Y)}(u,v) \tag{34}$$
$$\le \sup_V \pi(u|v) \le \pi_X(u) \;\Rightarrow\; \pi_X(u) = \sup_V \pi(u|v) \;.$$

Now, let $\phi(u) = \bigvee_v [\pi(u|v) \wedge \pi_Y(v)]$. Then

$$\text{(i)}\quad \pi(u|v) \wedge \pi_Y(v) \le \pi_X(u) \wedge \pi_Y(v) \tag{35}$$

since $\pi(u|v) \le \pi_X(u)$. Thus

$$\sup_V [\pi(u|v) \wedge \pi_Y(v)] \le \sup_v [\pi_X(u) \wedge \pi_Y(v)] = \pi_X(u)$$

and hence $\phi(u) \le \pi_X(u)$.

$$\text{(ii)}\quad \pi_X(u) \le \pi_Y(v) \;\Rightarrow\; \pi(u|v) \wedge \pi_Y(v) = \pi_{(X,Y)}(u,v) \wedge \pi_Y(v) \tag{36}$$
$$= \pi_{(X,Y)}(u,v)$$
$$\pi_X(u) > \pi_Y(v) \;\Rightarrow\; \pi(u|v) \wedge \pi_Y(v) = \pi_{(X,Y)}(u,v)\frac{\pi_X(u)}{\pi_Y(v)} \wedge \pi_Y(v)$$
$$\ge \pi_{(X,Y)}(u,v) \wedge \pi_Y(v)$$
$$\left(\text{since } \frac{\pi_X(u)}{\pi_Y(v)} > 1\right)$$
$$= \pi_{(X,Y)}(u,v)$$

and hence

$$\pi(u|v) \wedge \pi_Y(v) \ge \pi_{(X,Y)}(u,v) \;.$$

Now,

$$\phi(u) = \sup_v [\pi(u|v) \wedge \pi_Y(v)] \ge \sup_v \pi_{(X,Y)}(u,v) = \pi_X(u)$$

which implies that

$$\phi(u) \ge \pi_X(u)$$

and consequently

$$\pi_X(u) = \phi(u) = \sup_v [\pi(u|v) \wedge \pi_Y(v)] \;. \tag{37}$$

Remark. Given $\pi_{(X,Y)}$, if we associate with the variable Y a possibility distribution G which is not necessarily a marginal distribution that is induced by $\pi_{(X,Y)}$, we then say that $\pi_{(X,Y)}$ is particularized by G [1].

Denote by $\bar{G}$ the cylindrical extension of G, i.e.

$$\bar{G} = U \times G .$$

Then the max-min composition $\pi_{(X,Y)} \circ \bar{G}$ represents a particularized possibility distribution of X given that $\pi_Y = G$.

In particular,

$$\pi_X(u) = \bigvee_v [\pi_{(X,Y)}(u,v) \wedge \pi_Y(v)] \tag{38}$$

which shows that (33) holds also for non-normalized conditional possibility distributions. Thus the main reason for the normalization of the expression for a conditional distribution relates to the need for consistency with non-interaction.

4. Application

We proceed to outline how a problem in the analysis of a probabilistic system gives rise to an analogous problem involving possibility distributions.

Consider a time-invariant discrete-state system such as considered in [9]. In the notation of [9], when the system is probabilistic, from

$$\begin{aligned} x_i(\pi) &= E[C_{ij}(\pi_i) + x_j(\pi)] \\ &= \sum_{j=1}^{n+1} P_{ij}[C_{ij}(\pi_i) + x_j(\pi)] , \quad i = 1,\ldots,n \\ &= \sum_{j=1}^{n} P_{ij}(\pi_i)x_j(\pi) + \sum_{j=1}^{n+1} P_{ij}(\pi_i)C_{ij}(\pi_i) \end{aligned} \tag{39}$$

we obtain the vector equation

$$X(\pi) = P(\pi)X(\pi) + C(\pi) \ . \tag{40}$$

In [9] it is shown that if at least one proper policy exists, then the equation

$$X^0 = \underset{\pi}{\mathrm{Min}}[P(\pi)X^0 + C(\pi)] \tag{41}$$

has a unique solution which is the minimum expected cost vector associated with an optimal policy. It is given by $X^0 = \lim_{r\to\infty} T^r(X)$, where X is arbitrary;

$$T(x) = \underset{\pi}{\mathrm{Min}}[P(\pi)x + C(\pi)] \ ; \tag{42}$$

and T^r is the r^{th} iterate of T. The proof of this result involves an application of Banach's fixed point theorem.

Now, if the system is possibilistic, the notion of a conditional possibility distribution can be used to replace

$$P_{ij}(k) = \mathrm{Prob}[s_{t+1} = q_j | s_t = q_i,\ u_t = \alpha_k] \tag{43}$$

by a corresponding expression involving Poss in place of Prob. Then, the expectation is replaced by a weighted mean

$$X_i(\pi) = \frac{\sum_{j=1}^{n+1} \hat{P}_{ij}(\pi_i)[C_{ij}(\pi_i) + X_j(\pi)]}{\sum_{j=1}^{n+1} \hat{P}_{ij}(\pi_i)} \tag{44}$$

and

$$C_i(\pi) = \frac{\sum_{j=1}^{n+1} \hat{P}_{ij}(\pi_i)C_{ij}(\pi_i)}{\sum_{j=1}^{n+1} \hat{P}_{ij}(\pi_i)} \tag{45}$$

and thus

$$Y(\pi) = P(\pi)X(\pi) + C(\pi)$$

where $Y(\pi)$ is the vector $(\alpha_1 X_1(\pi),\ldots,\alpha_n X_n(\pi))$,

$$\alpha_i = \sum_{j=1}^{n+1} \hat{P}_{ij}(\pi_i)\ , \quad i = 1,2,\ldots,n\ , \tag{46}$$

and $P(\pi)$ is the $n \times n$ matrix with generic element $\hat{P}_{ij}(\pi_i)$, $\hat{P}_{ij}(\pi_i)$ being the transition possibility from q_i to q_j when command π_i is applied.

Define $S: \mathbf{R}^n \to \mathbf{R}^n$ by

$$s(X) = (S_1(X),\ldots,S_n(X))$$

where $S_i(X) = \frac{1}{\alpha_i}[(P\cdot X)_i + C_i]$. Now let $T(X) = \min_{\pi} S(X)(\pi)$. Then the optimal policy will correspond to the solution of

$$X^0 = T(X_0)\ . \tag{47}$$

5. Concluding Remarks

The concept of a possibility distribution serves to provide a basis for the analysis of situations in which the uncertainty is nonstatistical in nature. In addition, it is useful when the uncertainty is partly statistical and partly nonstatistical, as in the case where the probabilities are characterized in linguistic terms. In this context, the mathematical formulation described in this note provides a rigorous basis for defining a normalized conditional possibility distribution which may be used in a variety of situations. In a forthcoming paper, we shall describe additional applications of the concept of a conditional possibility distribution, particularly to the theory of belief and the mathematical theory of evidence (Shafer [4]).

Acknowledgment

The author wishes to acknowledge helpful discussions with Professor L.A. Zadeh relating to the concept of a conditional possibility distribution and its applications.

References

[1] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1977) 3-28.

[2] H.T. Nguyen, Some mathematical tools for linguistic probabilities, Proceedings of 1977 IEEE Symposium on Fuzzy Sets and Applications.

[3] R.E. Bellman and L.A. Zadeh, Local and fuzzy logics, in: J.M. Dunn and G. Epstein, eds., Modern Uses of Multiple-Valued Logic (1977) 103-165 (D. Reidel, Dordrecht).

[4] G. Shafer, A mathematical Theory of Evidence, Princeton University Press, Princeton (1976).

[5] J. Kampe de Feriet, Mesure de l'information fournie par un evenement, Coll. Inter. C.N.R.S. #186, Paris (1970) 191-221.

[6] H.T. Nguyen, Sur les mesures d'information de type inf., Lectures Notes in Mathematics, #398, Springer-Verlag (1974) 62-75.

[7] L.A. Zadeh, On fuzzy algorithms, Memo ERL-M325, University of California, Berkeley (1972).

[8] T.L. Fine, A mathematical theory of evidence, by G. Shafer, Book reviews, Bull. A.M.S. 83 (1977) 667-672.

[9] J.H. Eaton and L.A. Zadeh, Optimal pursuit strategies in discrete-state probabilistic systems, Journal of Basic Engineering 84 (1962) 23-29.

[10] N. Sugeno, Theory of fuzzy integrals and its applications, Tokyo Institute of Technology, Tokyo (1974).

[11] T. Terano and M. Sugeno, Conditional fuzzy measures and their applications, in: L.A. Zadeh, K.S. Fu, K. Tanaka and M. Shimura, eds., Fuzzy Sets and Their Applications to Cognitive and Decision Processes (Academic Press), New York, 1975.