SOME ASYMPTOTIC RESULTS FOR THE COMBINATION OF EVIDENCE PROBLEM

I.R. Goodman

Code 421
Information Systems Division
Naval Ocean Systems Center
San Diego, California 92152

Abstract.

The combination of evidence problem is treated here as the construction of a posterior possibility function (or probability function, as a special case) describing an unknown state parameter vector of interest. This function exhibits the appropriate components contributing to knowledge of the parameter, including conditions or inference rules, relating the parameter with observable characteristics or attributes, and errors or confidences of observed or reported data. Multivalued logic operators - in particular, disjunction, conjunction, and implication operators, where needed - are used to connect these components and structure the posterior function. Typically, these operators are well-defined for only a finite number of arguments. Yet, often in the problem at hand, a number of observable attributes represent probabilistic concepts in the form of probability density functions. This occurs, for example, for attributes representing ordinary numerical measurements- as opposed to those attributes representing linguistic-based information, where non-probabilistic possibility functions are used. Thus the problem of discretization of probabilistic attributes arises, where p.d.f.'s are truncated and discretized to probability functions. As the discretization process becomes finer and finer, intuitively the posterior function should better and better represent the information available. Hence, the basic question that arises is: what is the limiting behavior of the resulting posterior functions when the level of discretization becomes infinitely fine, and, in effect, the entire p.d.f.'s are used?

It is shown in this paper that under mild analytic conditions placed upon the relevant functions and operators involved, nontrivial limits in the above sense do exist and involve monotone transforms of statistical expectations of functions of random variables corresponding to the p.d.f.'s for the probabilistic attributes.

Keywords.

combination of evidence problem; knowledge-based systems; possibilities; multivalued logic; t-norms; t-conorms; asymptotic forms of t-conorms

INTRODUCTION

or the past few years there has been almost a litral explosion of interest in uncertainty modeling nd its applications to knowledge-based systems and the general field of Artificial Intelligence. See , for example: the survey of Bonissone and Tong (1985), the entire issue of the International Journal of Man-Machine Studies, Vol.22, of which it is a part; the survey of Bonissone and Brown(1985); the Proceedings of the 8^{th} (1983,at Karlsruhe) and 9^{th} (1985, at Los Angeles) International Joint Conferences on Artificial Intelligence; the Proceedings of the First Workshop on Uncertainty and Probability in Artificial Intelligence (Los Angeles, 1985); and the textbook of Goodman and Nguyen (1985).

Much of this activity is concerned with the ongoing controversies involving which approaches to uncertainty modeling are most appropriate and how to compare and contrast them. Among the more popular approaches, may be mentioned: subjective and objective probabilities; fuzzy sets and possibility functions with extensions of Zadeh's original theory; DS-Dempster-Shafer belief functions and related generalized measures; non-monotonic logics and more classical two-valued and multivalued logics; and various confidence and certainty factor techniques. (These and more may be found among the above references.)

The approach taken in this paper to uncertainty modeling can be considered a multivalued logical one, extending both classical probability and Zadeh's possibility approaches. (See Goodman and Nguyen (1985) for a complete exposition.) Because of these controversies, it is appropriate at this point to present some justification for the viewpont taken here.

Firstly, consider the concept of partial specification of probability distributions - or equivalently, random quantities. It is well-established doctrine to model a perceived situation from a probabilistic viewpoint, where various means, variances, higher order moments, or entire distributions are specified up to being in some classes. Determination of more specific values may take place later when empirical evidence warrants it, if the latter occurs at all. On the other hand, it has been shown (Goodman and Nguyen, 1985) that <u>all</u> possibility functions and large classes of associated multivalued logic or fuzzy set operators are equivalent, respectively, to the probabilistic concepts of specifying one point coverage probabilities for random sets and ordinary naturally corresponding set operators over random sets. It should be noted that random sets reduce to ordinary random variables or random vectors, when they are singleton-valued only; hence in those special cases, one point specification is the same as complete specification of random variables. Finally, one point coverage probabilities of random subsets of a given space can be shown to correspond to probabilities of interacting events having in common (through filter classes) given compound events represented as points in the space (Goodman, 1983). Consequently, the approach taken here can be considered a random set one with the rather weak specifications as described above. (See also Goodman (1984a) and Nguyen(1978) where DS is equated with r.sets)

Secondly, natural language descriptions are often intentionally vague - though not at all devoid of content - in character, representing compound, overlapping events with non-exact boundaries. (See e.g., Zadeh, 1981 for further comments in this direction. See Goodman and Nguyen, 1985, Chapter 2 for analysis of natural language through an initial formal language and category theory framework followed by semantic evaluations.) Thus, a weakened or underspecified probabilistic modelling approach is a plausible one for incorporating natural language information with more numerical/probabilistic type.

One immediate consequence of the equivalency of the approach to uncertainty modeling as presented in this paper and probabilities through specifications of random set coverage functions is a quantitative reduction of entropy for posterior estimation of an unknown parameter θ: When both linguistic-based data Y" and probabilistic, i.e., random, data Y' are used in the estimation, instead of only Y', the corresponding posterior entropy is lowered. Thus, in general,

$$\mathrm{Ent}(\theta|Y',Y") \overset{d}{=} \left(\underset{\substack{S(Y") \text{ arb in} \\ R(Y")}}{\mathrm{Ent}(\theta|Y',S(Y"))}\right) < \mathrm{Ent}(\theta|Y'), \quad (1)$$

where R(Y") is the class of all random sets which are one point equivalent to $\phi_{Y"}$, the possibility function representing Y", i.e.,

$$\phi_{Y"}(x) = \Pr(\text{ x in } S(Y")) , \quad (2)$$

for all x in the support space of $\phi_{Y"}$.

Another application of the equivalency of representation is to the admissibility-inadmissibility controversy concerning the comparison of classes of uncertainty measures. Lindley (1982), extending Savage's original argument, considered the choice of an uncertainty measure for a given situation as a decision theory problem, utilizing an additive type of loss function between uncertainty measure of possible events and the occurance-non-occurance of the events. He showed that, depending on the choice of particular loss function, only a fixed (non-identity in general) monotone function of any probability measure will be admissible (=Bayes). Thus, all other uncertainty measures are inadmissible within this context. However, this also includes all monotone transforms not of the fixed form on probabilities, including, in general, probability measures themselves! On the other hand, by first allowing admissible to include the union of all such admissible classes in the previous senses and by using the one point representation of any uncertainty measure (considered as a possibility function over some suitable collection of subsets such as a sigma algebra), the entire problem may be cast into deciding among classes of probability measures on a higher order space. In turn, with suitable modifications, it follows that all uncertainty measures will be admissible. (See Goodman and Nguyen, 1985, pp. 558-567 for details.)

COMBINATION OF EVIDENCE PROBLEM

The combination of evidence problem as usually defined includes such diverse situations as medical diagnoses, military assessment of combat and readiness states, detective work, using clues and experience, and control of complex sytems.

Suppose formally the following holds:

$\theta \overset{d}{=}$ unknown parameter state vector, in $\Omega \subseteq R^w$ (3)

$\theta_1 \overset{d}{=}$ joint vector of possible values of θ according to each of m sources relating to corresponding components of Z, in $\Omega^m \subseteq R^{wm}$ (4)

$Z \overset{d}{=}$ joint vector of characteristic or attribute values of potential observables connected with θ, in D (5)

$Y_1 \overset{d}{=}$ actual observation vector of Z in error (6)

$\theta_2 \overset{d}{=}$ joint vector of possible values of θ according to each of n direct sources, in $\Omega^n \subseteq R^{wn}$ (7)

$Y_2 \overset{d}{=}$ actual observation vector of θ_2 in error (8)

Then the un-normalized posterior possibility function describing

$$\underline{\theta} \overset{d}{=} (\theta_1,\theta_2) \quad (9)$$

through

$$\underline{Y} \overset{d}{=} (Y_1,Y_2) \quad (10)$$

is determined by the relation

$$\phi(\underline{\theta}|\underline{Y}) = \underset{(Z \text{ in } D)}{\phi_{or}}(\phi(\underline{\theta},Z|\underline{Y})) , \quad (11)$$

for some conjunction operator $\phi_{\&}$ such as a t-conorm or co-copula. (See discussion around (30)-(32).) In turn, the joint posterior is

$$\phi(\underline{\theta},Z|\underline{Y}) = \phi_{\&}(\phi(\theta_2|Y_2,\theta_1,Z,Y_1),\phi(\theta_1|Z,\underline{Y}), \phi(Z|\underline{Y})) , \quad (12)$$

for some conjunction operator $\phi_{\&}$ such as a t-norm or copula (Again, see the discussion around (30)-(32))

For simplicity, assume the sufficiency conditions

$$\phi(\theta_2|Y_2,\theta_1,Z,Y_1) = \phi(\theta_2|Y_2) , \quad (13)$$

$$\phi(\theta_1|Z,\underline{Y}) = \phi(\theta_1|Z). \quad (14)$$

Equations (9)-(14) hold for all $\underline{\theta}$,Z,$\underline{Y}$;in general, unless specified, all equations will be assumed to hold for all values of the relevant variables. Note the classical probability function analogues to the above decomposition of conditional forms with:

possibility function ϕ replacing prob.func. p ,
disjunction operator ϕ_{or} " " $\int(\cdot)d(\cdot)$,
conjunction operator $\phi_{\&}$ " " product.

Note also the equivalent alternative formulation analogue of Bayes' Theorem:

$$\phi_{\&}(\phi(\underline{\theta},Z|\underline{Y}),\phi(\underline{Y})) = \phi_{\&}(\phi(\underline{Y}|\underline{\theta},Z),\phi(Z|\underline{\theta}),\phi(\underline{\theta})) \overset{d}{=} \phi(\underline{Y},\underline{\theta},Z) , \quad (15)$$

$$\phi(\underline{Y}) = \underset{\left(\substack{Z \text{ in } D, \\ \underline{\theta} \text{ in } \Omega^{w(m+n)}}\right)}{\phi_{or}}(\phi(\underline{Y},\underline{\theta},Z)). \quad (16)$$

In the case in (15)-(16), the conditionals $\phi(\underline{Y}|\underline{\theta},Z)$, $\phi(Z|\underline{\theta})$ and prior $\phi(\underline{\theta})$ must be known in order to solve for $\phi(\underline{\theta},Z|\underline{Y})$ as given implicitly in (15). On the other hand, in (12)-(14), the three posteriors $\phi(\theta_2|Y_2)$,$\phi(\theta_1|Z)$,$\phi(Z|\underline{Y})$ must be known to solve for $\phi(\underline{\theta},Z|\underline{Y})$ as given explicitly in (12). In many knowledge-based systems, experts are used in conjunction with physical/ mathematical considerations so that the three posteriors can be obtained directly, where in particular, $\phi(\theta_1|Z)$ represents the joint inference rule effect, $\phi(Z|\underline{Y})$ represents the joint posterior confidences or errors between $\underline{Y}$ and Z, and $\phi(\theta_2|Y_2)$ may well be vacuous. (See, e.g., Goodman and Nguyen, 1985, Chapters 8,9.)

In order to determine the final normalized posterior possibility function for θ, the diagonal event must be introduced as

$$\mathcal{D} \overset{d}{=} \{\theta^{m+n}|\theta \text{ in } \Omega\} , \quad (17)$$

where the notation

$$\theta^{m+n} \overset{d}{=} (\theta,\theta,\ldots,\theta) \text{ (m+n terms)} \quad (18)$$

is used. It then follows that

$$\phi(\theta|\underline{Y}) \overset{d}{=} \phi(\underline{\theta}=\theta^{m+n}|\underline{Y}) = \phi(\underline{\theta}=\theta^{m+n},\mathcal{D}|\underline{Y}) = \phi_{\&}(\phi(\theta|\underline{Y},\mathcal{D}),\phi(\mathcal{D}|\underline{Y})), \quad (19)$$

where

$$\phi(\mathcal{D}|\underline{Y}) = \phi_{or}(\phi(\theta|\underline{Y})), \quad (\theta \text{ in } \Omega) \tag{20}$$

again using the basic properties of conditional possibility functions. Then solve for $\phi(\theta|Y,\mathcal{D})$ in (19) as given as an implicit function of $\phi(\theta|\underline{Y})$, as determined through (11), and $\phi(\mathcal{D}|\underline{Y})$ as given in eq. (20).

Throughout the above equations, some abuse of notation has occurred with respect to the logical operators $\phi_{\&}$ and ϕ_{or}. Although the same symbol may be used, it may refer to quite different operators of the same general type in different parts of the computations. For example, $\phi_{\&}$ may be a different conjunction operator in its use in (12) as compared with (19), or ϕ_{or} may not be the same disjunction operator in (20) as it is in (11), nor in the latter need the operator be the same throughout the entire range of values of Z . For a discussion of this issue and some guidelines for choosing logical operators, see Goodman (1984b).

SUBJECTIVE VS. PROBABILISTIC ATTRIBUTES

Consider now in more detail the joint vector of attributes associated with the unknown parameter. Some attributes can arise from probabilistic considerations, such as blood pressure readings, pulse measurements, and body temperature observations in medical diagnosis of a patient's condition. Other attributes may be concerned with more subjective things such as how the patient feels, how intense and where is the pain, and degree of difficulty in breathing. Similarly, in a military problem where a decision is to be made whether two track histories followed by two different sensor systems are the same or not, probabilistic attributes may well include geolocations and various sensor system parameter measurements, while subjective attributes could typically include verbal descriptions of the two targets involved.

Suppose for simplicity, from now on, each attribute contributing to the problem at hand is either probabilistic or subjective in form, so that Z can be partitioned as

$$Z = (Z', Z''), \tag{21}$$

where the superscript ( )' refers always to probabilistic attributes and ( )" refers to subjective ones. Thus one can write also for the domain of Z

$$D = D' \times D'' \tag{22}$$

and also assuming for simplicity from now on that θ_2 and Y_2 are vacuous so that in effect n = 0 ,

$$Y \overset{d}{=} \underline{Y} = Y_1 = (Y', Y'') , \tag{23}$$

and

$$\underline{\theta} = \theta_1 . \tag{24}$$

Typical probabilistic attributes arise from probability density functions from standard families such as multivariate gaussian or exponential. Some probabilistic attributes of the more finitely discrete type are naturally represented by probability functions. On the other hand, subjective attributes usually are determined through experts in the form of finite conditional confidence or error tables, using possibility or probability functions. Since possibility functions include probability functions as a subclass, it is also assumed from now on that Z' refers to only p.d.f.-derived attributes, while Z" refers to attributes characterized by finite numbers of domain values. But note that in general (ϕ_{or} = max being a notable exception) the operator (or operators-recall the above discussion of the last section) ϕ_{or} in (11) operates on only a finite number of arguments, i.e., D should be finite. Further, for purposes of implementation, only a finite number of operations should be used. In general, the form in (11) does not reduce to a more analytic structure and is indeed discrete in nature with Z varying arbitrarily in D, in order to be "integrated out" analogous to the situation for ordinary probability functions .

Consequently, for (11) to make sense, although D", and hence $\phi(Z''|Y'')$, may be left alone, D' must in effect be replaced by a suitably finitely discretized and truncated form, say D'_q , and $\phi(Z'|Y')$ likewise must be replaced or represented by a finitely discrete and truncated form $\phi_q(Z'|Y')$ of the original joint conditional p.d.f. f over $D' \subseteq \mathbb{R}^r$, say, where q is an index referring to the level of discretization and truncation. Thus, it is assumed

$$D'_1 \subset D'_2 \subset \cdots \subset D'_q \uparrow D' , \tag{25}$$

q = 1,2,3,..., in any natural sense of convergence, and for any Z' in D'_q ,

$$\phi_q(Z'|Y') = f(Z'|Y')\cdot\Delta_q(Z') , \tag{26}$$

where $\Delta_q(Z')$ is the r-volume of a corresponding small cell surrounding value Z' with edges dividing the nearest neighbors of Z' relative to D'_q . Thus, compatible with the above equations,

$$\Delta_1(Z') > \Delta_2(Z') > \cdots > \Delta_q(Z') \downarrow 0 . \tag{27}$$

With the above suppositions made, a natural question to ask is :

Is there some choice of class of logical operators such that probabilistic information can be utilized fully in obtaining the posterior function for θ, i.e., are

$$\phi_\infty(\theta|Y) \overset{d}{=} \lim_{q \to +\infty} (\phi_q(\theta|Y)) \tag{28}$$

and

$$\phi_\infty(\theta|Y,\mathcal{D}) \overset{d}{=} \lim_{q \to +\infty} (\phi_q(\theta|Y,\mathcal{D})) \tag{29}$$

well-defined ?

In the next sections it is shown that for a large class of t-conorms, with additional mild analytic conditions imposed upon the relevant operators and functions involved in computing the posteriors, the above question is answered in the affirmative. In addition, for disjunction operator max, though $\phi_\infty(\theta|Y)$ becomes zero, the normalized form $\phi_\infty(\theta|Y,\mathcal{D})$ is nontrivial. Specialization of the results is also carried out for a convenient subclass of operators defined originally by Frank(1979).

LIMITING FORMS OF POSTERIORS

Before obtaining the main results of this section, a brief review of an important class of disjunction operators will be given.

Disjunction operator $\phi_{or}:[0,1]^2 \to [0,1]$ is called an Archimedean t-conorm iff:it is a t-conorm- i.e., a nondecreasing, continuous, symmetric, associative function satisfying the usual boundary conditions

$$\max(x,y) \le \phi_{or}(x,y) \le \min(x+y,1), \tag{30}$$

$$\phi_{or}(x,0) = x \quad ; \phi_{or}(x,1) = 1 , \tag{31}$$

for all x,y in [0,1], and it is Archimedean, i.e.,

$$\phi_{or}(x,x) > x , \tag{32}$$

for all x with 0<x<1.

Note that, clearly, max is a non-Archimedean t-conorm. However, it can be shown that any t-conorm can, in a certain sense, be expressed as a superposition of affine transforms of max and Archimedean t-conorms. For this and other general results, see e.g. Goodman and Nguyen (1985, Chapter 2.3.6).

Note that any t-conorm can be extended in a natural recursive manner, unambiguously, from two to any finite number of arguments, by utilizing the associativity property. Dual remarks hold for t-norms and Archimedean t-norms, where: in eq.(30) max is replaced by min, $\le$ by $\ge$, min(x+y,1) by max(x+y-1,0); in (31) 0 is replaced by 1 and 1 by 0; in (32) > is replaced by <; and throughout, ϕ_{or} is replaced by $\phi_{\&}$.

The key property of Archimedean t-conorms needed here is the following theorem from the literature:

Theorem 1. (See Goodman and Nguyen, 1985, p. 116.)

1. If ϕ_{or} is any Archimedean t-conorm, then there is a generating function $h:[0,1] \to R^+ \cup \{+\infty\}$ with h being continuous, non-increasing, and

$$h(0) \le +\infty \quad ; \quad h(1) = 0 \ , \tag{33}$$

such that ϕ_{or} is the DeMorgan transform of some t-norm determined by h :

$$\phi_{or}(x_1,\dots,x_t)=1-\phi_{\&}\,(1-x_1,\dots,1-x_t) \ , \tag{34}$$

for all $x_1,\dots,x_t$ in [0,1] and

$$\phi_{\&}(y_1,\dots,y_t)=h^{-1}(\min(h(0),\ \sum_{k=1}^{t} h(y_k))), \tag{35}$$

for all $y_1,\dots,y_t$ in [0,1]; all t = 1,2,3,... . (36)

2. Conversely, for any choice of generating function h given in (33) and $\phi_{\&}$ constructed as in (35) and, in turn, ϕ_{or} determined as in (34), ϕ_{or} will be an Archimedean t-conorm. ∎

Consider now the following assumptions and notation to be used in the chief results:

(I) The following order of operations is assumed in (11) at any discretization-truncation level q:

$$\phi_{or}(\cdot) = \underset{(Z \text{ in } D_q)}{\phi_{or''}}(\ \underset{(Z'' \text{in } D'')}{\ }\ \underset{(Z' \text{ in } D_q')}{\phi_{or'}}(\cdot)\ ), \tag{37}$$

where in general, $\phi_{or'}$ and $\phi_{or''}$ are distinct disjunction operators.

(II) $\phi_{or'}$ is any fixed Archimedean t-conorm (not depending upon index q) with generator function h such that $d^2h(x)/d^2x$ is continuous in x for all $1-\varepsilon_1 \le x \le 1$, for some $0 < \varepsilon_1 < 1$.

(III) Referring only to $\phi_{\&}$ as used in (12) and assuming it is a t-norm, it is also supposed that $\partial^2\phi_{\&}(x,y)/\partial^2y$ is continuous in y for any given x and uniformly bounded in x and y, for all $0\le x\le 1$ and $0 \le y \le \varepsilon_2$, for some $0 < \varepsilon_2 < 1$.

(IV) Referring to eq.(26), f is assumed to be uniformly bounded over D' by finite constant M, for all possible Y'.

(V) The following order of operations is assumed in (11) at any discretization-truncation level q:

$$\phi_{or}(\cdot) = \underset{(Z \text{ in } D_q)}{\ }\ \underset{(Z' \text{ in } D_q')}{\phi_{or'}}(\ \underset{(Z'' \text{ in } D'')}{\phi_{or''}}(\cdot)\ ), \tag{38}$$

where in general, $\phi_{or'}$ and $\phi_{or''}$ are distinct disjunction operators.

(VI) $\phi_{or''}$ considered as a function over $[0,1]^{D''}$ has all of its second order (including mixed) derivatives continuous over $[0,\varepsilon_3]^{a_o}$, where a_o denotes the number of elements of D".

Consider also the following definitions:

$$\psi(x) \overset{d}{=} 1-h^{-1}(\min(h(0),x)) \ , \tag{39}$$

for all x in $R^+ \cup \{+\infty\}$;

$$\omega(x) \overset{d}{=} (\partial\, \phi_{\&}(x,y)/\partial y)_{y=0} \ , \tag{40}$$

for all x in [0,1], where $\phi_{\&}$ refers to (12);

$$\mu(x) \overset{d}{=} \beta_0\cdot\omega(x) \ , \tag{41}$$

for all x in [0,1] ;

$$\phi(\theta,Z''|Y'',Z') \overset{d}{=} \phi_{\&}(\phi(\theta|Z),\phi(Z''|Y'')) \ ; \tag{42}$$

$$\kappa_\infty(\theta,Z''|Y) \overset{d}{=} E_{(Z'|Y')}(\mu(\phi(\theta,Z''|Y'',Z'))), \tag{43}$$

where formally (Z'|Y') is considered a conditional random vector with p.d.f. f as given in (26) ;

$$\beta_0 \overset{d}{=} -(d\ h(x)/dx)_{x=1} \ ; \tag{44}$$

$$\alpha_0 \overset{d}{=} (\partial\Big(\underset{(Z'' \text{ in } D'')}{\phi_{or''}}(v(Z''))\Big)/\partial v(Z''))_{V=0}, \tag{45}$$

where

$$V \overset{d}{=} (v(Z''))_{Z'' \text{ in } D''} \ , \tag{46}$$

with

$$0 \le v(Z'') \le 1 \tag{47}$$

allowed to be arbitrary ;

$$n(\theta,Z'|Y'') \overset{d}{=} \alpha_0\cdot \sum_{(Z'' \text{ in } D'')} (\mu(\phi(\theta,Z''|Y'',Z'))); \tag{48}$$

$$\nu_\infty(\theta|Y) \overset{d}{=} E_{(Z'|Y')}(n(\theta,Z'|Y'')), \tag{49}$$

where formally (Z'|Y') is considered a conditional random vector with p.d.f. f as given in (26).

With all of the preliminaries established, the chief result follows. This shows that, up to certain non-creasing transforms involved, the limiting posteriors in (28),(29) are determined through expectations.

Suppose, as usual, $\phi_q(\theta|Y)$ and $\phi_q(\theta|Y,\mathcal{D})$ denote the un-normalized and normalized, respectively, posterior possibility functions for θ at discretization-truncation level q, as presented in the previous section.

Theorem 2. (Modification and extension of Goodman and Nguyen, 1985, Chp. 9(F).)

1. Suppose assumptions (I),(II),(III),(IV) all hold. Then

$$\phi_\infty(\theta|Y) = \underset{(Z'' \text{ in } D'')}{\phi_{or''}}(\psi(\kappa_\infty(\theta,Z''|Y))) \ , \tag{50}$$

where ψ is a non-decreasing function with $\psi(0)$=0, given in (39) and where κ_∞ is the expectation of non-decreasing function μ, with $\mu(0)$=0, given in (41) (through ω given in (40) and non-negative constant β_0 given in (44)), where the argument of μ is the quantity $\phi(\theta,Z''|Y'',Z')$, a function of (Z'|Y') considered here formally as a conditional random variable.

2. Suppose assumptions (II),(III),(IV),(V),(VI) all hold. Then

$$\phi_\infty(\theta|Y) = \psi(\nu_\infty(\theta|Y)) \ , \tag{51}$$

where ψ (non-decreasing, etc.) is given in (39) and ν_∞ is the expectation of n , as given in (48) considered as a function of formal conditional random variable (Z'|Y').

Proof 1:

Applying Theorem 1 to $\phi_{or'}$ from assumption (II),

$$\underset{(Z' \text{ in } D_q')}{\phi_{or'}}(\phi_q(\theta,Z|Y)) = \psi(\kappa_q(\theta,Z''|Y)), \tag{52}$$

where

$$\kappa_q(\theta,Z''|Y) \overset{d}{=} \sum_{(Z' \text{ in } D_q')} (h(1-\phi_q(\theta,Z|Y))). \tag{53}$$

Now by assumption (II) again,

$$\xi(x,y) \overset{d}{=} h(1-\phi_{\&}(x,y)) \tag{54}$$

$$= \mu(x)\cdot y \ + R(x,y) \ , \tag{55}$$

where remainder $R(x,y)$ satisfies

$$|R(x,y)| \le (y^2/2)\cdot B_\varepsilon \ , \qquad (56)$$

for some $B_\varepsilon > 0$, all y in $[0,\varepsilon]$, for some fixed ε, $1 > \varepsilon > 0$, and all x , $1 \ge x \ge 0$. Then substituting into (55),(56)

$$x = x(Z') \stackrel{d}{=} \phi(\theta,Z''|Y'',Z') \ , \qquad (57)$$

$$y = y(Z') \stackrel{d}{=} \phi_q(Z'|Y') \ , \qquad (58)$$

holding arbitrary but fixed θ,Z'',Y, and noting that by choosing q so large that

$$\mathrm{card}(D_q') > M\cdot c_0/\varepsilon \ , \qquad (59)$$

it follows that for all Z' in D_q' , $y(Z') < \varepsilon$, where constant c_0 is such that

$$c_0 \ge \mathrm{card}(D_q') \cdot \max_{(Z' \text{ in } D_q')}(\Delta_q(Z')), \qquad (60)$$

yields the summations

$$Q_q \stackrel{d}{=} \sum_{(Z' \text{ in } D_q')} (\xi(x(Z'),y(Z'))) = T_q + W_q \qquad (61)$$

where

$$T_q \stackrel{d}{=} \sum_{(Z' \text{ in } D_q')} (\mu(x(Z')),y(Z')) \qquad (62)$$

and

$$W_q \stackrel{d}{=} \sum_{(Z' \text{ in } D_q')} (R(x(Z'),y(Z'))). \qquad (63)$$

Now, from (56),(59),(60), it follows that

$$\lim_{q \to +\infty} (W_q) = 0 \qquad (64)$$

and hence from (61),(54),(53)

$$\lim_{q \to +\infty} (\kappa_q(\theta,Z''|Y) = \lim_{q \to +\infty} (T_q) = \int_{(Z' \text{ in } R^r)} \mu(x(Z'))\cdot f(Z'|Y')dZ'. \qquad (65)$$

Proof 2:

The proof is analogous to that for part 1, except that first one must expand the function $h(1-\lambda(\zeta(Z',y)))$ in y in some sufficiently small closed interval $[0,\varepsilon]$ about $y=0$, where

$$\lambda(V) \stackrel{d}{=} \phi_{or''} \underset{(Z'' \text{ in } D'')}{(v(Z''))} \ , \qquad (66)$$

V as in (46),(47), and where

$$\zeta(Z',y) \stackrel{d}{=} (\zeta(Z'',Z',y))_{Z'' \text{ in } D''}, \qquad (67)$$

$$\zeta(Z'',Z',y) \stackrel{d}{=} \phi_\&(x(Z',Z''),y), \qquad (68)$$

utilizing (57),(58).

■

SPECIFIC LIMITING FORMS

It is desirable to determine as large a family as possible of t-norms and t-conorms which will satisfy all the assumptions required for Theorem 2. One such family is due originally to Frank (1979) who characterized the complete solution of all t-norms and t-conorms such that the modularity condition

$$\phi_{or}(x,y) = x+y-\phi_\&(x,y) \ ; \text{ all } x,y \text{ in } [0,1]. \qquad (69)$$

The Archimedean solution is given in terms of parameter s , $0<s\le\infty$, and because $s=1$, $s=+\infty$ and the special case non-Archimedean solution $s=0$ are all limiting cases, denote G for the set of all $s, 0<s<+\infty$, $s \ne 1$. Also, in the following results, all arguments $x,y,x_1,\ldots,x_t$ etc. are assumed to lie in $[0,1]$:

$$\phi_{or,s}(x,y) = x+y-\phi_{\&,s}(x,y) = 1-\phi_\&(1-x,1-y) \qquad (70)$$

and for multiple arguments, more generally,

$$\phi_{or,s}(x_1,\ldots,x_t) = 1-\phi_{\&,s}(1-x_1,\ldots,1-x_t), \qquad (71)$$

with generator function h_s given as

$$h_s(x) = -\log((s^x-1)/(s-1)), \ s \text{ in } G, \qquad (72a)$$

and the limiting cases

$$h_s(x) = \begin{cases} -\log(x) \ , & s=1 \\ 1-x \ , & s=+\infty \\ \delta(x) \text{ (Dirac delta)}, & s\to 0 \end{cases} \qquad (72b)$$

yielding

$$\phi_{\&,s}(x_1,\ldots,x_t) = \begin{cases} \log_s(1+ (\prod_{j=1}^{t}(s^{x_j}-1)/(s-1)^{t-1})), & s \text{ in } G \\ x_1\cdots x_t \text{ (product)} & s=1 \\ \max(0, \sum_{j=1}^{t} x_j -(t-1)) \ , & s=+\infty \\ \min(x_1,\ldots x_t) \ , & s=0 \end{cases} \qquad (73)$$

and

$$\phi_{or,s}(x_1,\ldots,x_t) = \begin{cases} 1-\log_s(1+(\prod_{j=1}^{t}(s^{1-x_j}-1)/(s-1)^{t-1})), & s \text{ in } G \\ 1- \prod_{j=1}^{t} (1-x_j) \text{ (prob. sum)} \ , & s=1 \\ \min(1, \sum_{j=1}^{t} x_j)\text{(bnded sum)} \ , & s=+\infty \\ \max(x_1,\ldots,x_t) \ , & s=0 \ . \end{cases} \qquad (74)$$

It follows that assumptions (II),(III),(VI) all hold here for all $s>0$. More specifically,

$$0>dh_s(x)/dx = \begin{cases} -s^x \log(s) / (s^x-1) \ , & s \text{ in } G \\ -1/x \ , & s=1 \\ -1 \ , & s=+\infty \\ -\delta(x) \ , & s\to 0 \ . \end{cases} \qquad (75)$$

Hence,

$$0 < \beta_{0,s} = \begin{cases} s\log s \ / \ (s-1) \ , & s \text{ in } G \\ 1 \ , & s=1, +\infty \\ -\infty \ , & s\to 0 \ , \end{cases} \qquad (76)$$

$$0\le d^2h_s(x)/d^2x = \begin{cases} (\log(s))^2s^x/(s^x-1)^2 \ , & s \text{ in } G \\ 1/x^2 \ , & s=1 \\ 0 \ , & s=+\infty \\ \delta(x) \ , & s\to 0 \ , \end{cases} \qquad (77)$$

and thus h_s is convex and hence $\phi_{\&,s}$ is a copula (see Goodman and Nguyen, 1985,Chp. 2.3.6) . Also,

$$0\le\partial\phi_{\&,s}(x,y)/\partial y = \begin{cases} (s^x-1)s^y/(s-1+((s^x-1)(s^y-1))), & s \text{ in } G \\ x \ , & s=1 \\ \left.\begin{matrix} 0 \ , & x+y<1 \\ 1 \ , & x+y>1 \end{matrix}\right\} \ , & s=+\infty \\ \left.\begin{matrix} 0 \ , & x<y \\ 1 \ , & x>y \end{matrix}\right\} \ , & s=0 \ , \end{cases} \qquad (78)$$

$$0\le\partial^2\phi_{\&,s}(x,y)/\partial^2 y = \begin{cases} (s^x-1)s^y\log(s)\cdot(s-s^x)/(s-1+((s^x-1)(s^y-1))^2, & s \text{ in } G \\ 0 \ , & s=1,+\infty,0 \end{cases}$$

$$\le \begin{cases} s\cdot\log(s) \ , & s\ge 1 \\ |\log(s)|/s^2 \ , & 0<s<1. \end{cases} \qquad (79)$$

Eqs.(39),(40),(72),(78) imply

$$\psi_s(x) = \begin{cases} 1-\log_s(1+((s-1)\cdot e^{-x})), & s \text{ in } G \\ 1-e^{-x} \ , & s=1 \\ x \ , & s=+\infty \\ 1-\delta_{x,0} \text{ (Krönecker delta)}, & s\to 0 \ , \end{cases} \qquad (80)$$

$$\omega_s(x) = \begin{cases} (s^x-1)/(s-1) \ , & s \text{ in } G \\ x \ , & s=1 \\ 0 \ , & s=+\infty \\ 1-\delta_{x,0} \ , & s\to 0 \ , \end{cases} \qquad (81)$$

$$0<\partial\phi_{or,s}(x_1,\ldots,x_t)/\partial x_j = \begin{cases} \prod_{1\le i\le t,\ i\ne j}(s^{1-x_i}-1)\cdot s^{1-x_j}/((s-1)^{t-1}+ \prod_{1\le i\le t}(s^{1-x_i}-1)) \ , & s \text{ in } G \\ \prod_{1\le i\le t,\ i\ne j}(1-x_i) \ , & s = 1 \ , \end{cases} \qquad (82)$$

etc., and hence from (45),(82),

$$\alpha_{0,s} = 1 \ , \ 0<s\le+\infty \ . \qquad (83)$$

Finally, it can be shown after some calculations

$$|\partial^2\phi_{or,s}(X)/\partial x_j\partial x_k| \le \begin{cases} s^2\log(s)/(s-1) \ , & +\infty>s>1 \ , j\ne k \\ s\cdot\log(s) & , j=k \ , \end{cases} \qquad (84a)$$

$$|\partial^2\phi_{or,s}(X)/\partial x_j \partial x_k| \le \begin{cases} |\log(s)|/(s^2\cdot(1-s)), & 0<s<1, \; j\ne k \\ |\log(s)|/s^2 & ,\; j=k \\ 0 \;,\; s=1 \text{ or } s=+\infty . \end{cases} \tag{84b}$$

Thus in summary:

Corollary 1.

1. If assumptions (I),(IV) are made and Frank's family of Archimedean operators is chosen to be used in all of the computations for $\phi_q(\theta|Y)$, then the conclusions of part 1 of Theorem 2 hold, using (76),(80),(81).

2. If assumptions (IV),(V) are made and Frank's family of Archimedean operators is chosen to be used in all of the computations for $\phi(\theta|Y)$, then the conclusions of part 2 of Theorem 2 hold, using (76),(80),(81),(83).

∎

Note that from (19), the explicit solution using $\phi_{\&,s}$

$$\phi_q(\theta|Y,\mathcal{D})=\log_s\left(1+((s^{\phi_q(\theta|Y)}-1)(s-1)/(s^{\phi_q(\mathcal{D}|Y)}-1))\right), \tag{85}$$

for all $s > 0$ and all q, including $q=+\infty$.

Noting that for the non-Archimedean operators min and max (corresponding to $s=0$ in Frank's family of operators), at least condition (II) is violated (via (75),(77)), so that Theorem 2 is not applicable here. However, a modification of the theorem can be carried out as is given next.

Theorem 3.

Suppose now

$$\phi_{or} = \phi_{or'} = \phi_{or''} = \max \tag{86}$$

with assumptions (III),(IV) also holding. Suppose finally in (26) that independent of any Z' in D'_q

$$\Delta_q(Z') \equiv \Delta'_q . \tag{87}$$

Then

1.
$$\phi_\infty(\theta|Y) \equiv 0 . \tag{88}$$

2. But, nontrivially,

$$\phi_\infty(\theta|Y,\mathcal{D}) = \omega^{-1}(\rho(\theta,Y)/\rho(Y)) , \tag{89}$$

where

$$\rho(\theta,Y) \overset{d}{=} \max_{(Z'' \text{ in } D'', Z' \text{ in } R^r)} (\omega(\phi(\theta,Z''|Y'',Z'))\cdot f(Z'|Y')), \tag{90}$$

$$\rho(Y) \overset{d}{=} \max_{(\theta \text{ in } \Omega)} (\rho(\theta,Y)) = \max_{(Z'' \text{in } D'', Z' \text{ in } R^r)} (\omega(g(Z,Y''))\cdot f(Z'|Y')), \tag{91}$$

where

$$g(Z,Y'') \overset{d}{=} \phi_\&(\phi(\theta=\theta^*(Z)|Z), \phi(Z''|Y'')) , \tag{92}$$

with $\theta^*(Z)$ defined through

$$\phi(\theta=\theta^*(Z)|Z) = \max_{(\theta \text{ in } \Omega)}(\phi(\theta|Z)). \tag{93}$$

Proof:

First consider here (11),(12),(20). Then apply the expansion for small x, Δ'_q with remainder $\bar{R}$

$$\phi_\&(x,y\cdot\Delta'_q)=\omega(x)\cdot y\cdot\Delta'_q + \bar{R}(x,y\cdot\Delta'_q) \tag{94}$$

and in turn substitute the resulting expressions for $\phi_q(\theta|Y)$ and $\phi_q(\mathcal{D}|Y)$ into (19) and again use (94) and then solve for $\phi_q(\theta|Y,\mathcal{D})$. Finally, take limits as $q\to+\infty$, showing the remainders go to 0. ∎

Example: Probabilistic-Only Matching of a Vs. b.

Suppose one wishes to estimate θ, the degree of association between objects a and b, where: no subjective components are present, i.e., $D''=\emptyset$, etc.; all observed data is stat. indep. 2-dimensional gaussian; a single inference rule is used with exponential intensifiers/modifiers ϕ_α for antecedent and neutral matching function ϕ_A determined from use of of the standard weighted statistical distance test statistic τ for testing the null hypothesis equality of means $H_0: u_a=u_b$; and Frank's family is used: (95)

$$f(Z'|Y') \overset{d}{=} f_a(Z'_a|Y'_a)f_b(Z'_b|Y'_b), \; f_i(Z_i|Y_i) \overset{d}{=} N_2(u_i(Y'_i),\Lambda_i),$$

$$\phi(\theta|Z)=\phi_s(\theta|Z') \overset{d}{=} \phi_{\Rightarrow,s}(\phi_{ant}(Z'),\phi_{cons}(\theta)), \tag{96}$$

$$\phi_{\Rightarrow,s}(x,y) \overset{d}{=} \phi_{or,s}(\phi_{nt}(x),y), \; \phi_{nt}(x) \overset{d}{=} 1-x, \; \phi_{ant}(Z') \overset{d}{=} \phi_\alpha(\phi_A(Z')), \tag{97}$$

$$\phi_\alpha(x) \overset{d}{=} x^\alpha \;,\; \alpha \ge 0 \;;\; \phi_\alpha(x) \overset{d}{=} (1-x)^{|\alpha|},\; \alpha < 0 , \tag{98}$$

$$\phi_A(Z')=Pr(\tau(X)>\tau(Z')|H_0)=Pr(\chi^2_2>\tau(Z'))=e^{-\tau(Z')/2}, \tag{99}$$

$$\tau(Z') \overset{d}{=} (Z'_a-Z'_b)^T\cdot(\Lambda_a+\Lambda_b)^{-1}\cdot(Z'_a-Z'_b), \text{etc.} \tag{100}$$

Then (101)

$$\phi_\infty(\theta|Y')=\psi_s(E_{(Z'|Y')}(\mu_s(\phi(\theta|Z')))=\psi_s(\beta_{0,s}\cdot\delta_s(\theta|Y')),$$

$$\delta_s(\theta|Y') \overset{d}{=} E_{(Z'|Y')}(\omega_s(\phi_s(\theta|Z'))) , \tag{102}$$

$$\omega_s(\phi_{\Rightarrow,s}(x,y)) = \begin{cases} s/(s-1+((s^x-1)(s^{1-y}-1))), & s \text{ in } G, \\ 1-((1-y)\cdot x) \;,\; s=1 . \end{cases} \tag{103}$$

$$\delta_s(\theta|Y')=E_{\chi^2_{2,\tau(Y')}}\left(\omega_s(\phi_{\Rightarrow,s}(\phi_\alpha(\exp-\chi^2_{2,\tau(Y')}/2),\phi_{cons}(\theta)))\right) \tag{104}$$

since r.v.$(\tau(Z')|Y')$ is $\chi^2_{2,\tau(Y')}$. Assuming $\phi_{cons}(\theta_c)=1$, for at least some θ_c in Ω, using $\phi_{or}=\max$ in (20), and using (85), (105)

$$\phi_\infty(\mathcal{D}|Y')=\psi_s(\beta_{0,s})=1-\log_s(1+((s-1)\cdot s^{(s-1)/s})), \tag{106}$$

$$\phi_\infty(\theta|Y',\mathcal{D}) = \log_s\left(1+\frac{(s^{\phi_\infty(\theta|Y')}-1)(s-1)}{s^{1+(1/((s-1)s^{(s-1)/s}))}}\right).$$

In particular for the limiting case $s=1$, (107)

$$\phi_\infty(\mathcal{D}|Y')=1-e^{-1};\; \phi_\infty(\theta|Y',\mathcal{D})=(1-\exp-\delta_1(\theta|Y'))/(1-e^{-1}),$$

$$\delta_1(\theta|Y')=1-((1-\phi_{cons}(\theta))g(\alpha,\tau(Y'))), \tag{108}$$

$$g(\alpha,\tau(Y'))=E_{\chi^2_{2,\tau(Y')}}(\phi_\alpha(\exp-(\chi^2_{2,\tau(Y')}/2))) \tag{109}$$

$$= \begin{cases} (1/(\alpha+1))\cdot\exp-((\alpha/(2(\alpha+1)))\cdot\tau(Y')) \;,\; \alpha\ge 0 \\ \sum_{k=0}^{\infty} ((-1)^k|\alpha|^{[k]}/(k+1)!)\cdot\exp-((\tau(Y')/2)\cdot(k/(k+1))), \end{cases} \tag{110}$$

where

$$x^{[k]} \overset{d}{=} \begin{cases} x\cdot(x-1)\cdots(x-(k-1)) \;,\; k=1,2,\ldots \\ 1 \;,\; k=0 \;, \end{cases} \tag{111}$$

using the Poisson mixture form of noncentral chi-square r.v.'s. (See, e.g., Johnson and Kotz, 1972.)

Acknowledgments

This work has been supported by the NOSC IR/IED Program (Indep.Research and Indep.Explor.Develop.).

REFERENCES

Bonissone, P.P. and A.L. Brown (1985). Expanding the horizons of expert systems. Proc. Conf. Expt. Sys. & Know.Eng., Ruschlikon, Zurich, Switz, Gottlieb Duttweiler Inst. North-Holland, N.Y.

Bonissone, P.P. and R.M. Tong (1985). Editorial: reasoning with uncertainty in expert systems. Int.J. Man-Machine Stud., 22, 241-250.

Frank, M.J. (1979). On the simultaneous associativity of F(x,y) and x+y-F(x,y). Aeq.Math., 19, 194-226.

Goodman, I.R. (1983). A unified approach to modeling and combining of evidence through random set theory. Proc. 6 MIT/ONR Wrkshp. C^3 Sys., 42-47.

Goodman, I.R. (1984a). Some new results concerning random sets and fuzzy sets. Inf.Sci., 34, 93-113

Goodman, I.R. (1984b). An approach to the target data association problem using subjective and statistical information. Proc. 1984 Amer.Cntrl. Conf., 587-592.

Goodman, I.R. and H.T. Nguyen (1985). Uncertainty Models for Knowledge-Based Systems. North-Holland, Amsterdam.

Johnson, N. and S. Kotz (1972). Distributions in Statistics, Vol. 3, Chap.28, John Wiley & Sons, N.Y.

Lindley, D.V. (1982). Scoring rules and the inevitability of probability. Int.Stat.Rev., 50, 1-26.

Nguyen, H.T. (1978). On random sets and belief functions. J.Math.Anal.& Applic., 65, 531-542.

Zadeh, L.A. (1981). Test score semantics for natural languages and meaning representation via PRUF. In B.B. Rieger (Ed.), Empirical Semantics. Bochum, Brockmeyer, pp. 281-349.