Conditions for the Existence of Belief Functions Corresponding to Intervals of Belief

by

John F. Lemmer
CTA Incorporated
200 Liberty Plaza
Rome, NY 13440
jlemmer@nova.npac.syr.edu

Henry E. Kyburg, Jr.
Department of Computer Science
University of Rochester
Rochester, NY 14627
Kyburg@cs.rochester.edu

Abstract

While every Shafer belief function corresponds to a set of interval beliefs on the atoms of the frame of discernment, an arbitrarily specified set of intervals of belief may not correspond to *any* belief function, even when it does correspond to bounds imposed by sets of probability functions. This paper proves necessary and sufficient conditions which must be met by a set of belief intervals over atoms if a corresponding belief function exists. The sufficiency is proved via an an O(n) algorithm which will always construct an corresponding belief function, if one exits, for a specific set of intervals

Introduction[1]

C. A. B. Smith (Smith 1961), I. J. Good (Good 1962), and H. E. Kyburg (Kyburg 1961) all early on had the idea that the representation of belief should be by means of intervals rather than by real numbers. There are intuitive reasons for this, of course, but only C. A. B. Smith offered a pragmatic argument. While it is true, as Savage (Savage 1954) argued, that one can be *forced* to choose among alternatives in such a way as to reveal in the imaginary limit a real-valued degree of belief, at some point the choices will strike the agent making them as arbitrary, forced by the demands of the interlocutor rather than by the agent's belief state. Smith's idea was that degrees of belief should correspond to what he called "pignic" odds. The pignic odds on S are the least odds that the agent would feel comfortable in offering on S.

It is clear that in general the probabilities corresponding to the pignic odds on S and to the pignic odds on the negation of S will not add up to one. For example, if I will offer odds of 1:2 on rain tomorrow, and odds of 1:1 on no rain tomorrow, the corresponding probabilities are 1/3 for rain and 1/2 for no rain. Since offering odds of less than 1:3 on no rain is like receiving odds of greater than 3:1 on rain, the agent's beliefs concerning rain can be represented by the probability interval for S : [1/3,1/2](The characterization of *evidential probability* developed about the same time (Kyburg 1961), in which all probabilities are based on imperfectly known statistical proportions, also leads to an interval representation of uncertainty.)

The analogy, so far, with Shafer's belief functions (Shafer 1976) is very close. It has been shown (Kyburg 1987) that every belief function corresponds to a convex set of classical probability functions. Smith's pignic odds correspond to suspending judgment on a convex set of classical probabilities. A.P Dempster discussed the relation between upper and lower probabilities and convex sets of probability functions in (Dempster 1967).

The problem of ensuring that a set of real-valued probabilities over an algebra are consistent (coherent) is non-trivial. The problem of ensuring the consistency of a set of pignic odds or a set of belief functions is even harder. By consistency of a set of probability intervals, we mean that there exists at least one classic probability function which takes a point from each interval. The assignment [.1,.2], [.3,.4] to exclusive and exhaustive S_1, S_2 is obviously not consistent; while the assignment [.1,.9], [.4,.5] is since the probability function can assign a value of .5 to $P(S_1)$ and .5 to $P(S_2)$. If this same set of intervals is to be consistent in the sense of a belief function, a belief function must exist which yields these intervals as the belief and upper probability of these singletons.

In this paper we offer an algorithm for determining the consistency of a set of interval constraints on atoms of an

[1] The first author acknowledges the support of New York State Science and Technology Foundation. The second author acknowledges the support of the Signals Warfare Center of the U. S. Army and of the National Science Foundation.

event space relative to the existence of a belief function, and for generating one of the many belief functions that correspond to that set of constraints, in case the constraints are consistent. The algorithm yields a complete belief (or mass) function.

In applications, the initial interval constraints can be provided by an expert. Many people are more comfortable with interval estimates than with real-valued estimates. When probabilities are obtained by formal statistical methods, they are often given by confidence intervals. If we are to use interval inputs in any form, some procedure for checking their consistency is crucial. While our program does this for only a limited special case, there is hope that it can be extended. The algorithm and a convenient man-machine interface have been implemented (Lemmer, 1988).

Background

This paper investigates the use of belief intervals as 'input' to the Dempster–Shafer calculus. Many implemented AI systems using models of uncertainty based on Dempster-Shafer Theory rely on intervals to elicit the belief functions, e.g. (Garvey, Lowrance, & Fischler 1981). But using intervals of belief in this way confronts the difficulty that many sets of intervals of belief do not correspond to *any* belief function. This is so even for intervals that are probabilistically consistent; see (Kyburg 1987).

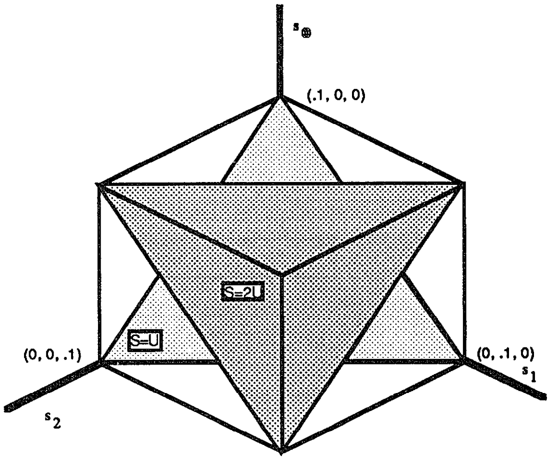

Figure 1: Regions of Existence

In the three event case shown in Figure 1 and described in the next section, only one fifth of all the probabilistically consistent intervals are consistent with any belief function. his situation can prove quite frustrating to a domain expert building a knowledge based system when he is forced by the theory to revise his original estimates. It is especially frustrating if he has no sure guidance on how to carry out this revision.

Correspondence Between Interval Probability Constraints and Belief Functions

It is well known that for every belief function Bel, Bel(X) and P*(X) = 1 − Bel(~ X) can be construed as the minimum and maximum probabilities of X under a convex set of probability functions (Kyburg 1987). Such an interval can be computed for each event in the frame of discernment on the basis of the belief function Bel. But for what subjective intervals of belief can a belief function be consistently posited? That is the subject of this section. Note that since the intervals cannot be specified arbitrarily, neither can belief in the complement of an event be specified arbitrarily.

The criteria which must be met by a set of interval constraints on atoms in order for that set to correspond to a belief function are mathematically straight-forward. Assume that we have a frame of discernment $\theta = \{e_0, e_1, \cdots, e_{n-1}, \}$, where the members of θ are ordered by the size of the belief interval associated with them. Let $[l_i, u_i]$ be the interval associated with e_i, and let $i < j$ imply $(u_i - l_i) \le (u_j - l_j)$.

The size of the interval associated with e_i is $s_i = u_i - l_i$, and the sum of these sizes is

$$S = \sum_{0}^{n-1} s_i$$

The belief uncommitted to any singleton atom of θ, U, is

$$U = 1.0 - \sum_{0}^{n-1} l_i$$

Theorem 1.
There exists a belief function corresponding to a set of intervals over atoms (or a set of basic subsets) of a frame of discernment if and only if

$$U \ge 0 \qquad (1)$$

$$\forall i \, [s_i \le U] \qquad (2)$$

$$S \ge 2U \qquad (3)$$

The necessity of (1), (2), and (3) follows almost directly from the definition of a belief function. The sufficiency of (1), (2), and (3) will be shown by presenting an algorithm which will always construct a valid basic probability assignment whenever the three conditions are met.

Before actually providing the proof that (1), (2), and (3) are necessary and sufficient, it is worthwhile examining their implication. Suppose that the frame of discernment, θ, is the set $\{e_0, e_1, \cdots, e_{n-1}, \}$, and the set of probability constraints on these events is

$$\begin{Bmatrix} 0.2 \le e_0 \le 0.2+s_0 \\ 0.3 \le e_1 \le 0.3+s_1 \\ 0.4 \le e_2 \le 0.4+s_2 \end{Bmatrix}$$

What do the conditions, tell us about the values which the s_i can assume and still have a belief function exist satisfying these constraints? The answer in the case of these constraints is shown in Figure 1. Using the formula for U introduced above, we can compute the value of U in the example as 0.1. Constraints (2) and (3) imply that the values for the s_i must lie in the tetrahedron between the darker gray plane shown in the figure, and the point (.1,.1,.1): Constraint (2) limits possible solutions to the cube. Constraint (3) limits them to the positive side of the darker shaded plane. Probability intervals are consistent within the cube on positive side of the lighter shaded plane.

Necessity and Sufficiency

The necessity of (1), (2), and (3) follows almost directly from the definition of a belief function. The sufficiency of (1), (2), and (3) will be shown by presenting an algorithm which will always construct a valid basic probability assignment whenever the three conditions are met.

A belief function exists if and only if (Shafer 1976) there exists a basic probability assignment, $m: 2^{\Theta} \rightarrow [0,1]$, such that

$$m(\varnothing)=0 \tag{4}$$

$$\sum_{A \subset \Theta} m(A) = 1.0 \tag{5}$$

and

$$Bel(A) = \sum_{B \subset A} m(B) \tag{6}$$

From (4), (6) and the definition of l_i, it follows that

$$Bel(\{e_i\}) = m(\{e_i\}) = l_i \tag{7}$$

and from the definition of U, (5), and (7) that

$$U = 1.0 - \sum_i l_i = 1.0 - \sum_i m(\{e_i\}) = \sum_{\substack{A \subset \Theta \\ |A| \ge 2}} m(A) \ge 0$$

because all the m(.) are non-negative. Thus we have the necessity of condition (1).

Since

$$s_i = P^*(\{e_i\}) - Bel(\{e_i\}) = \sum_{\substack{A \subset \Theta \\ e_i \in A}} m(A) - m(\{e_i\}) = \sum_{\substack{A \subset \Theta \\ |A| \ge 2 \\ e_i \in A}} m(A) \le \sum_{\substack{A \subset \Theta \\ |A| \ge 2}} m(A) = U$$

we have the necessity of condition (2).

Because

$$S = \sum_i s_i = \sum_i \sum_{\substack{A \subset \Theta \\ |A| \ge 2 \\ e_i \in A}} m(A) = \sum_{\substack{A \subset \Theta \\ |A| \ge 2}} |A|\, m(A) \ge 2 \sum_{\substack{A \subset \Theta \\ |A| \ge 2}} m(A) = 2\left[1.0 - \sum_{e \in \Theta} m(\{e\})\right] = 2U$$

we have the necessity of condition (3).

The algorithm for constructing a belief function meeting a set of constraints satisfying conditions (1), (2), and (3) operates in four phases. The first phase begins with an (improper) basic probability assignment in which $m(\varnothing) = 1.0$, and all other assignments are zero. At the conclusion of Phase I, the (still improper) basic probability assignment has been modified so that, if the belief assigned to the empty set is excluded from the sum in (6), the belief calculated for each event, Bel $(\{e_i\})$, yields the correct value for l_i . At the end of phase two, computation of $P^*(e_i)$ (again excluding the empty set)will yield u_i for all the events in the (ordered) frame of discernment except the event with the largest sized interval, e_{n-1}. At the end of the third phase computation of $P^*(e_{n-1})$ (still with the same exclusion) will yield the correct result. At the end of the fourth phase, $m(\varnothing)$ will be zero, and all constraints will be satisfied by the (now proper) basic probability assignment.

Phase I: *Satisfy belief constraints*:

Set $m(\varnothing)$ equal to 1.0;
For all i, $0 \le i < n$,
  Set $m(\{e_i\})$ equal to l_i;
  decrease $m(\varnothing)$ by l_i;

It is clear that at the conclusion of this phase the value of $m(\varnothing)$ will always be equal to U. However, this value will be incrementally decreased in each of the following phases until it becomes 0 when the algorithm successfully terminates. Since $m(\varnothing)$ will become 0 and the remaining phases will never alter the values of the e_i, the condition (8) holds when the algorithm

$$l_i = \sum_{\substack{A \subset \Theta \\ A \subset \{e_i\}}} m(A) = m(\{e_i\}),\ 0 \le i < n \qquad (8)$$

successfully terminates.

In the remaining phases, $m(\varnothing)$ will decrease in such a way that if it ever becomes less than 0 we will be able to conclude that the value of s_{n-1} (and possibly others) is greater than U, violating condition (2); if termination occurs with $m(\varnothing)$ greater than 0 we will be able to conclude that condition (3) is violated.

Phase II: *Satisfy the upper probability constraints for all events except e_{n-1}. At the beginning of each iteration, U^* is the value of s_i which could be computed from the current state of the belief function (excluding the value of $m(\varnothing)$).*

Set R equal to Θ and U^* equal to 0;
for all i, $0 \le i < n-1$:
  Set $m(R)$ equal to $s_i - U^*$;
  set U^* equal to $U^* + m(R)$;
  decrease $m(\varnothing)$ by $m(R)$;
  remove e_i from R;
  if $m(\varnothing)$ is less than 0, s_i is greater than U.

At the conclusion of this phase, the correct value for the upper probability of each element of the frame of discernment, except for the one with the largest interval, can be computed from the belief function so far constructed, that is :

$$\forall i,\ 0 \le i < n-1 \left[u_i = \sum_{\substack{B \subset \Theta \\ \{e_i\} \cap B \ne \varnothing}} m(B) \right] \qquad (9)$$

Moreover,

$$U = m(\varnothing) + U^* \qquad (10)$$

The first condition, (9), is guaranteed by the construction of R and the first operation in the 'for' loop. The second condition, (10) is guaranteed by the operations on U^* and $m(\varnothing)$. From the second condition it can be seen that if $m(\varnothing)$ has a value less than zero then the interpretation of U^* implies that s_i must be greater than U. This implies that one of the necessary conditions, i.e. (2), was violated by the constraint set.

At the termination of this phase U^* is also the the value which could be computed for s_{n-1}, neglecting the value of $m(\varnothing)$. This is because every set, R, for which m(R) became non-zero during Phase II contained e_{n-1} (and at least one more atom)..

The essential action of the next phase is to preserve conditions of (9) and (10) while satisfying the upper probability constraint for e_{n-1}.

Phase III: *Satisfy the upper probability constraint for event e_{n-1}. In what follows the interpretation of U^* changes from the previous phase: it now refers only to the amount of s_{n-1} so far accounted for, without any reference to R. S is simply a temporary set which always contains e_{n-1} and one other event.*

Set R equal to Θ;
for all i, $0 \le i < n-2$:
  if U^* is equal to s_{n-1} terminate this phase;
  Set S equal to $\{e_i, e_{n-1}\}$;
  set $m(S)$ equal to $\min\{[s_{n-1} - U^*], m(R)\}$
  decrease $m(R)$ by $m(S)$;
  decrease $m(\varnothing)$ by $m(S)$;
  if $m(\varnothing) < 0$ s_{n-1} is greater than U;
  increase U^* by $m(S)$;
  remove e_i from R;
  increase $m(R)$ by $m(S)$;
if U^* is not equal to s_{n-1} then s_{n-1} is greater than U.

For Phase III to terminate successfully, U* must equal s_{n-1}, and $m(\varnothing)$ be greater than or equal to 0. We now show failure to satisfy either of these conditions implies that the constraints did not satisfy the necessary conditions.

First note that conditions (9) and (10) remain valid throughout the execution of this phase: (10) trivially so since in each iteration whatever is subtracted from $m(\varnothing)$ is added to to U^*.

The preservation of condition (9) is less obvious. Subtracting m(S) from m(R) reduces s_j for all j such that e_j is an element of R. However these values are restored for all these j's except j=i by removing e_i from R and adding to the new m(R) the value of m(S). The value of s_i had been raised by adding value to m(S) but is restored by the subtraction of value from m(R) while R still contains e_i.

The net effect of each iteration is to raise U^* (the amount of s_{n-1} accounted for), reduce $m(\emptyset)$, and preserve (9) and (10).

If $m(\emptyset)$ becomes negative, (10) and the meaning of U^* imply immediately that s_{n-1} is greater than U, violating (2).

If U^* is not equal to s_{n-1} when this phase terminates, the min operation will have guaranteed that U^* must be less than s_{n-1}. Once all iterations of the 'for' loop have been completed, all of U^* will have been assigned to m(A) such that every A contains e_{n-1} and *exactly* one other variable. Thus we will have that

$$s_{n-1} > U^* = \sum_{i=0}^{n-2} s_i$$

Because

$$S = s_{n-1} + \sum_{i=0}^{n-2} s_i$$

the inequality expression allows us to write

$$S = s_{n-1} + (1-x)\, s_{n-1}$$

where x must be positive so that

$$(1-x)\, s_{n-1} = U^* = \sum_{i=0}^{n-2} s_i$$

By necessary condition (3) we can write

$$S = s_{n-1} + (1-x)\, s_{n-1} \geq 2\,U$$
$$s_{n-1} \geq \left(\frac{2}{2-x}\right) U$$
$$s_{n-1} > U$$

which means that the situation we are analyzing cannot arise. Thus when Phase III terminates U^* must equal s_{n-1}. and condition (9) is subsumed by

$$\forall i, 0 \leq i < n \left[u_i = \sum_{\substack{B \subset \Theta \\ \{e_i\} \cap B \neq \emptyset}} m(B) \right] \qquad (11)$$

All we must do now is preserve (11) while insuring that $m(\emptyset)$ takes on the value of zero.

<u>Phase IV:</u> *Convert m to a proper basic probability assignment.*

Set R equal to Θ;
for all i, $0 \leq i < n-2$:
  Set x equal to
$$\min\left\{ m(R), \frac{|R| - 1}{\binom{|R|}{2} - |R| + 1} m(\emptyset) \right\};$$
  decrease m(R) by x;
  for $\forall\, t\, [t \subset R;\ |t| = 2]$:
    Increase m(t) by $\frac{1}{|R| - 1}$ x;
    decrease $m(\emptyset)$ by $\left[\frac{1}{|R| - 1} - \frac{1}{\binom{|R|}{2}} \right]$ x;
   if $m(\emptyset)$ equals 0 then terminate.
   Remove e_i from R;
if $m(\emptyset)$ does not equal 0 then S < 2U.

We need to show three things to claim that the final phase works correctly: that (11) is preserved, that $m(\emptyset)$ does not become negative, and that, if the phase terminates with positive $m(\emptyset)$, then S was less than twice U. These arguments go through without difficulty.

To see that (11) is preserved, note than when x is subtracted from m(R) the represented (in m) value of s_j for every j such that e_j is in R is reduced by x. When $\frac{x}{|R| - 1}$ is added to each of the $|R| - 1$ cardinality two subsets of R containing e_j, the originally represented value is restored.

To see that $m(\emptyset)$ never becomes negative begin by noting that the total added to the belief of the subsets of R will be greater than the amount subtracted from the belief in R. This is true because the amount $\frac{x}{|R| - 1}$ will be added $\binom{|R|}{2}$ times to subsets, and $\binom{|R|}{2} \frac{1}{|R| - 1}$ is greater than 1 for all |R| greater than 2. Indeed the reason we are doing this phase is to take this additional belief away from $m(\emptyset)$ while preserving (11). The most we want to take away is the amount of $m(\emptyset)$ itself. Thus x can be no more than that defined by the following expression:

$$x + m(\emptyset) = \binom{|R|}{2} \frac{x}{|R| - 1}$$

which can be solved for x giving

$$x = \frac{|R| - 1}{\binom{|R|}{2} - |R| + 1} m(\varnothing)$$

But m(R) cannot become negative either, hence the minimum function.

If the phase terminates with positive $m(\varnothing)$, then all iterations of Phase IV must have been performed. Therefore, all the belief mass which was in $m(\varnothing)$ at the end of Phase I(i.e a mass equal to U), is now either still in $m(\varnothing)$ or is in some of the m(A) such that the cardinality of A is two. This implies that

$$S = 2 \sum_{\substack{A \subset \Theta \\ |A|=2}} m(A)$$

and that

$$U = m(\varnothing) + \sum_{\substack{A \subset \Theta \\ |A|=2}} m(A)$$

allowing us to conclude that S < 2U, which is not allowed by condition (3).

Therefore we can conclude that conditions (1), (2), and (3) are sufficient to allow the construction of a belief function, because the algorithm just presented will always construct one when these conditions are met.

Future Work

The first important generalization of these results would be the construction of an algorithm that would take interval assessments of any set of subsets of the event space, rather than just the singleton subsets or a set of basic sets, ensure consistency, and derive a belief function for the whole space. Second, one would like to characterize, not just one, but the whole (infinite!) set of belief functions consistent with a set of consistent interval constraints. A beginning of such a characterization is available in (Lemmer 88). Third, since the representation of uncertainty by belief functions is a special case of the representation of uncertainty by *sets*` of probability functions over the event space, and there are indeed sensible situations in which belief functions won't do (Kyburg:1987), we would like to generalize these results to apply to the representation of uncertainty by sets of probability functions whether or not they correspond to belief functions.

References

Dempster, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping, *The Annals of Mathematical Statistics* 38: (325-339)

Garvey, T.D., Lowrance, J.D., and Fischler, M.A. 1981. An Inference Technique for Integrating Knowledge from Disparate Sources. In Proceedings of the Seventh IJCAI Vol 1:319-325

Good, I.J., Subjective Probability As a Measure of a Non-Measureable Set, *Logic, Methodology, and Philosophy of Science*. Nagel, E. Suppes, P. and Tarski, A. (eds) University of California Press, 1962:319-329

Kyburg, H.E. 1962. *Probability and the Logic of Rational Belief*. Wesleyan University Press

Kyburg, H.E., 1987. "Bayesian and Non-Bayesian Evidential Updating". *Artificial Intelligence* 31: (271-294)

Lemmer, J.F. 1988. Causal Probabalistic Reasoning System, Final Technical Report SBIR (88)-160, Knowledge Systems Concepts, Inc, Rome NY

Savage, L.J. 1954. *Foundations of Statistics.* John Wiley and Sons

Shafer, G. 1976 *A Mathematical Theory of Evidence.* Princeton University Press.

Smith, C.A.B. 1961. Consistency in Statistical Inference and Decision. *Journal of the Royal Statistical Society Series B"* (23):1-37