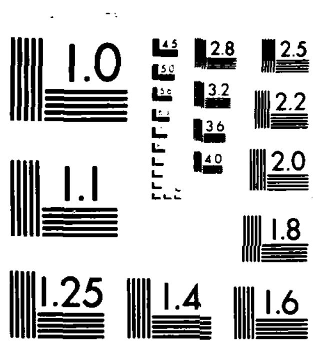
AD-A147 526 BELIEF STRUCTURES(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER M GOLDSTEIN AUG 84 MRC-TSR-2738 DAAG29-80-C-0041 1/1

UNCLASSIFIED F/G 12/1 NL

END
DATE
FILMED
12-84
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

MRC Technical Summary Report #2738

BELIEF STRUCTURES

Michael Goldstein

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

August 1984

(Received August 1, 1984)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

BELIEF STRUCTURES

Michael Goldstein*

Technical Summary Report #2738
August 1984

ABSTRACT

A belief structure is an inner product space which expresses one's (probabilistic) beliefs about certain random quantities. This structure is a fundamental element of the subjectivist theory of probability. This report provides an essentially non-technical description of the basis of the construction, and the ways in which it differs from more familiar Bayesian constructions.

AMS (MOS) Subject Classifications: 62A15

Key Words: Belief; coherence; prevision; projection.

Work Unit Number 4 - Statistics and Probability

* University of Hull, Department of Mathematics, Hull, England HU6 7RX.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

BELIEF STRUCTURES

Michael Goldstein*

(1) Introduction

"In investigating the reasonableness of our own modes of thought and behaviour under uncertainty, all we require, and all that we are reasonably entitled to, is consistency among these beliefs, and their reasonable relation to any kind of relevent objective data ("relevent" in as much as subjectively deemed to be so). This is Probability Theory ....

This point of view is not bound up with any particular philosophical position, nor is it incompatible with any such. It is strictly reductionist in a methodological sense, in order to avoid becoming embroiled in philosophical controversy."

(De Finetti, 1974. Preface.)

When Bayesian methods were first introduced into statistics, the emphasis was on the validity of using "subjective" rather than "objective" probabilities. As the main concern was for this broad general issue, the precise details of the formulation of the Bayesian theory have always remained somewhat unclear. Now that the Bayesian approach has become part of the "establishment" of ideas, it is time to subject the theory to proper critical scrutiny. Our intention is to rebuild the foundations of the subjectivist theory from first principles, essentially from the viewpoint expressed in the above quotation.

* University of Hull, Department of Mathematics, Hull, England HU6 7RX.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

What we shall argue is that subjectivist theory has a much broader range of implication than is expressed within the conventional Bayesian paradigm, and that many of the logical and practical difficulties in the Bayesian approach disappear when Bayesian theory is viewed as a special case of the more general theory. (This does not magically make hard problems become easy, but it does place the difficulty where it should be, namely in the questions that you are trying to answer, rather than in inappropriate methods for trying to answer these questions.)

Let us emphasize that we study the "reasonableness of our modes of thought" as a subject of interest in its own right. That is, precisely because the subject of probability finds application in so many diverse fields, it is important to expend the effort to ensure that the theory itself is sufficiently well-constructed to be worth applying. The danger is that, because of its immediate usefulness, probability theory can be, and often is, built around a particular methodology to which it appears well-suited. The theory is then confused with the methodology, and is "frozen" into a rigid form which is inappropriate for other applications. Indeed, once we confuse the "reasonableness of our modes of thought" with the "convenience" of certain modes of thought in certain applications, then it is impossible to distinguish between what is necessary, what is reasonable and what is arbitrary in our reasoning.

This article, while complete in itself, is intended as the first step in the systematic reformulation of the subjectivist theory. We will describe a construction, which we term a "belief structure", which we claim to be the proper setting for the theory. This assertion can only be demonstrated by providing a complete account of the theory developed in such a manner that our construction does play the fundamental role. This account will be the subject of a further series of articles, developing material which first appeared in Goldstein (1981, 1983, 1984, a,b,c). All that we intend in the present article is to describe the belief structure in sufficient detail to clarify the basis of the construction and the ways in which it differs from more familiar Bayesian constructions. As the belief structure depends on the properties of prevision, as developed by de Finetti, we begin with a brief discussion of this quantity.

(2) Prevision

The most difficult problem in constructing the subjectivist theory lies in distinguishing between aspects of the theory that are truly fundamental and aspects which are simply special cases, but seem fundamental because of convention, or because they appear well-suited to a particular application.

The first problem of this nature comes when we try to identify the basic expression of belief which we should consider. The choice that we face is whether to make probability or expectation the fundamental quantity in the theory. Most people are not aware of the importance of this choice. Perhaps this is because this issue is not important in traditional statistics (except at a purely technical level), and this attitude is subconsciously transferred to the subjectivist field. Further the visible example of the subjectivist approach is the Bayesian paradigm is which probability is pre-eminent.

However, such an expectation based methodology is possible and indeed necessary to overcome various logical and practical difficulties which arise in the Bayesian paradigm. We will address these difficulties in later articles. For now, let us simply recognise that we have a choice to make, and examine the way in which expectation can be formalized as a fundamental notion.

We will follow in this subsection the development of de Finetti (1974, 1975). De Finetti argues for expectation as the basis of theory, shows how this concept may be operationally realized and derives the basic properties of expectation from considerations of coherence. De Finetti's treatment is the basic motivation for our subsequent development.

De Finetti offers the following intuitive interpretation for expectation. Suppose that X is a random quantity (i.e. a numerical quantity which under some well-defined circumstance will be revealed but whose value is, at present, not known to you). Consider a ticket which will entitle you to a cash payoff of precisely X money units, when the value of X is revealed. Call the "price" of X, $P^*(X)$, the fair price (as judged by you) for the ticket on X, i.e. so that you are indifferent between the random gain X and the certain gain $P^*(X)$.

There are two intuitive properties of "price". Firstly, you clearly wouldn't pay more than the maximum payoff on the ticket, and presumably you would pay at least the minimum payoff on the ticket, i.e.

$$\inf X \le P^*(X) \le \sup X.$$

Secondly, if you imagine going into a "ticket shop" and buying two tickets, one on X and one on Y, then together you will have bought a ticket on $(X + Y)$, but you will have paid $P^*(X) + P^*(Y)$, suggesting that

$$P^*(X) + P^*(Y) = P^*(X + Y).$$

Thus $P^*(X)$ intuitively has the properties that we usually associate with expectation. Of course, this argument is entirely heuristic, and simply suggests a relationship, rather than offering a precise argument. You need not consider the "fair price" for a combination of articles to be the sum of the individual prices, and we have not demonstrated any undesireable consequences of deviating from such an assessment. However, de Finetti shows that by suitably modifying the definition of price we can precisely realize the intuitive notion we have introduced.

Thus, the intuitive notion of the price of a random gain X is formalized into the notion of the prevision of X as follows.

Definition If X is a random quantity, then your prevision of X, $P(X)$, is the value $\bar{x}$ which you would choose if, having made this choice, you were to suffer a penalty, or loss, L given by

$$L = K(X - \bar{x})^2$$

where K is a suitable constant defining the units of "loss".

(The above definition looks rather different from the fair price definition that we gave above. You can trace the informal relationship between the two approaches by expanding L, and demonstrating (heuristically) that you prefer the penalty $\frac{1}{2}(\bar{x} + c)$ to the penalty X if and only if $\bar{x} > c$.)

Suppose that you express previsions $\bar{x}_1,\ldots,\bar{x}_K$ for random quantities $X_1,\ldots,X_K$. By what criteria may you judge the reasonableness of your assessments? De Finetti proposes the following coherence criterion.

Criterion Your assessments $\bar{x}_1,\ldots,\bar{x}_K$ are said to be coherent if there is no other choice $a_1,\ldots,a_K$ with the property that for each possible outcome $X_1 = x_1,\ldots,X_K = x_K$ we have

$$\sum_i K_i(X_i - a_i)^2 < \sum_i K_i(X_i - \bar{x}_i)^2 .$$

In other words, you do not have a preference for a given (overall) penalty if you have the option of choosing an alternative penalty which will certainly be smaller, whatever the outcome.

This criterion seems reasonable. The only aspect of the outcome X_i which is relevent to your choice $\bar{x}_i$ is the numerical value of the associated penalty $K_i(X_i - \bar{x}_i)^2$, so that the only relevent feature of the combined outcome $(X_1,\ldots,X_K) = (x_1,\ldots,x_K)$ is the total penalty $\sum_i K_i(X_i - \bar{x}_i)^2$ Thus, all that we require is that if one penalty is certainly smaller than another penalty, then you should prefer the smaller penalty.

Using the above criterion, de Finetti provides the following necessary and sufficient conditions for a set of previsions $P(X_i) = \bar{x}_i$, $i = 1,\ldots,n$ to be coherent.

<u>Theorem</u> A specification $P(X_i) = \bar{x}_i$, $i = 1,\ldots,n$ is coherent if and only if the point $(\bar{x}_1,\ldots,\bar{x}_n)$ lies in the closed convex hull of the set of possible values of the random vector $(X_1,\ldots,X_n)$ in n-dimensional Euclidean space.

The above theorem contains, as immediate corollaries, the results

$$\inf(X) \leq P(X) \leq \sup(X)$$

$$P(a_1X_1 + \ldots + a_KX_K) = a_1P(X_1) + \ldots + a_KP(X_K)$$

We may use similar arguments to demonstrate a related property of prevision, namely that penalty A is preferred to penalty B if and only if $P(A) < P(B)$.

The above definition of prevision determines a quantity which plays the role usually taken by "expectation". However, while the two concepts are closely related, they are also fundamentally different. Your prevision for X is defined directly as a primitive quantity, while your expectation for X is a formal mathematical quantity defined in terms of your prespecified probability distribution over the possible values of X.

In this development, the probability of an event A is simply $P(A)$, where A is now also used as the indicator function of the event A (i.e.

A = 1 if "A occurs", A = 0 otherwise). There is no difference, in this formulation, between probability and expectation, and this is reflected in the common notation for the two concepts. Of course, this is quite separate from the psychological issue as to what classes of random quantities you personally prefer to specify previsions for. However, what we gain from this formulation is that we may consider "average" properties of random quantitites without exhaustively considering all possible outcomes. Thus you may specify all the probabilities in simple problems, but when you come to problems involving a large number of random quantities, you need not misrepresent your actual assessments by pretending that you have carefully assigned a many-dimensional probability distribution.

(3) Linear Structure

We have discussed the way in which you may make individual expressions of belief (i.e. previsions). However, theory says very little about how you should make such individual statements. Rather, it addresses the relationships between these statements. Thus, the object of study in the theory is not the individual belief statement, but rather a collection of belief statements, and one of the central concerns of the theory is the arrangement of such a collection in a manner which makes explicit the constraints imposed by coherence.

In this article, we shall describe one such arrangement, which we shall term a "belief structure". Our subsequent study of the foundations will turn out to be a study of the properties of belief structures. All the properties that we usually require of the subjective theory will turn out to have their fundamental expression within this structure. A closely related construction is described in de Finetti (1974, section 4.17). De Finetti briefly exhibits the construction to give simple geometric illustrations of general ideas. As we view the structure as fundamental, our notation, definitions and motivation are rather different from that of de Finetti.

The structure that we require is built in two stages. The first step is to make the linear structure of prevision explicit. Thus, suppose that you have specified your prevision for each of a collection $X_1,\ldots,X_n$ of random quantities. This also specifies your previsions for all linear combinations $\sum_{i=1}^{n} a_i X_i$ (and in general does not identify your prevision for any other function of $X_1,\ldots,X_n$). Thus the collection of linear combinations of $X_1,\ldots,X_n$ is a basic object of interest. We add the unit constant, X_0(i.e. $X_0 \equiv 1$), and denote by $\mathcal{L}$ the collection of all linear combinations of X_0, $X_1,\ldots,X_n$. $\mathcal{L}$ is a vector space in which each X_i is represented as a vector, and linear combinations of vectors are the corresponding linear combinations of random quantitites. Remember that, for example, X and X^2 are different random quantities (as they are not linearly related, unless X is an indicator function) even though they depend on the same (mathematical) variable. Thus, they will correspond to different vectors X_i, X_j (provided that you declare previsions for both X and X^2). It is up to you to explicitly decide which functions of which quantities to include as vectors in $\mathcal{L}$.

Within $\mathcal{L}$, we have the basic decomposition, for each element $X \in \mathcal{L}$, that

$$X = P(X)X_0 + (X-P(X)X_0)$$

In other words, we have decomposed each X into the sum of two vectors. The first, $P(X)X_0$, is known with certainty. The second, $(X - P(X)X_0)$, is "fair" (i.e. has "fair price", or prevision, zero).

The second stage of our construction, involves the addition of sufficient "geometry" over L so that the above decomposition corresponds to a basic structural property.

(4) Geometric Structure

Let us rephrase the definition of prevision in terms of the vector space L. The prevision of the vector X is the choice of vector $\bar{x}X_0$ to minimize the penalty $K(X - \bar{x}X_0)^2$. Preferring penalty A to penalty B corresponds to $P(A) < P(B)$. Thus, $P(X)X_0$ has the property that

$$P(X - P(X)X_0)^2 < P(X - cX_0)^2$$

for any $c \neq P(X)$.

Thus, if we measure "distances" between vectors by the "norm",

$$\|X - Y\| = \left(P(X - Y)^2\right)^{1/2}$$

corresponding to the inner product

$$(X,Y) = P(XY),$$

then the vector $P(X)X_0$ is the nearest vector to X of the form cX_0, i.e. the nearest element of the "plane of certainty" to X, i.e. the orthogonal projection of X into the subspace of L spanned by the single vector X_0.

This immediately implies that $(X - P(X)X_0)$ is orthogonal to X_0, so that

$$0 = (X_0, X - P(X)X_0) = P(X_0(X - P(X)X_0)) = P(X - P(X)X_0),$$

i.e. the vector $X - P(X)X_0$ must be fair. Thus, the decomposition outlined in section (3) can be expressed, via the inner product as the orthogonal decomposotion of X into two terms, one lying in the plane of certainty and one orthogonal to that plane.

Linearity is the property which underlies all subjectivist theory. What we shall demonstrate in subsequent articles is that the correspondence between prevision and projection, under this inner product, is the basic link which enables us to develop the theory in a manner in which the fundamental dependence on linearity (and nothing else) remains at all times explicit. Thus, we will formally define the inner product structure that we shall require.

(5) Belief Structure

Our results concerning the ways in which your beliefs change will, for the most part, be expressed in terms of the changes in your inner product, as introduced in the previous section. Thus, let us give this inner product space a name and formal definition. We will refer to it as a *belief structure*, and it is defined as follows.

Definition A belief structure is an inner product space, A, constructed as follows.

(i) Choose a specified collection of random quantitites, $C = X_0, X_1, \ldots, X_n$ (always including the unit constant, X_0), where n may be finite or infinite.

(ii) Construct the linear space, L, of all finite linear combinations of elements of C.

(iii) Construct the inner product and norm over L defined for each $X, Y \in L$ by

$$(X,Y) = P(XY), \quad \|X\| = (X,X)^{1/2} .$$

(We restrict C to contain elements X for which $P(X^2) < \infty$.)

Notice that $(\cdot,\cdot)$ defines a valid inner product precisely because of the properties of prevision, namely

(i) prevision is linear, so that

$$\begin{aligned}(X_1 + X_2, Y) &= P((X_1 + X_2)Y) \\ &= P(X_1Y) + P(X_2Y) \\ &= (X_1,Y) + (X_2,Y)\end{aligned}$$

(ii) prevision lies within the range of possible values, so that

$$(X,X) = P(X^2) > 0$$

(unless X "coincides with 0", i.e. you express a prevision of zero for every event of form $|X| > \varepsilon$, as $P(X^2) \geq \varepsilon^2 P(|X| > \varepsilon)$, for any ε. Strictly, we should identify equivalence classes of vectors whose differences "coincide" with 0.)

In the space $\mathcal{A}$, a "fair" vector, Y, is one which is orthogonal to X_0 $\left(\text{as } (X_0,Y) = P(Y)\right)$. Call $\mathcal{A}_0$ the subspace of $\mathcal{A}$ spanned by X_0. Fair vectors are vectors orthogonal to $\mathcal{A}_0$. Thus, the property that $P(X)X_0$ is the projection of X into $\mathcal{A}_0$ follows immediately as $(X-P(X)X_0)$ is a fair vector. (This is a simpler argument than that given previously, but it does not relate so directly to the definition of prevision.)

Note that we can add two belief structures $\mathcal{B}$ and $\mathcal{D}$, by constructing the combined inner product space, denoted $\mathcal{B} + \mathcal{D}$, which is spanned by all the vectors in $\mathcal{B}$ and $\mathcal{D}$. This will involve making the additional set of previsions $P(bd)$ for each $b \in \mathcal{B}$, $d \in \mathcal{D}$, to extend the inner product to the whole space. An important special case is where $P(bd) = 0$ for all $b \in \mathcal{B}$, $d \in \mathcal{D}$, so that in the combined space $\mathcal{B} + \mathcal{D}$, $\mathcal{B}$ and $\mathcal{D}$ are orthogonal subspaces, written $\mathcal{B} \perp \mathcal{D}$. We denote the orthogonal sum by $\mathcal{B} \oplus \mathcal{D}$.

The term "belief structure" reflects the two distinct aspects of A, (i) the purely structural aspects, as reflected in the choice of elements of C, (ii) the beliefs, as reflected in the inner product. Beyond this, the term is intended to reflect the following concept.

When you make individual prevision statements, these reveal certain individual aspects of your beliefs. However, individual prevision statements do not, of themselves, carry a very interesting crop of consequences about your beliefs. It is only when we examine a collection of prevision statements that systematic conclusions may be drawn. Further, certain collections of prevision statements will lead to interesting conclusions, while other collections of previsions may carry no particular implications of interest.

Thus, practical understanding of the theory does not usually proceed by simply making an arbitrary collection of prevision statements and then examining them for possibly interesting consequences. Rather, you begin with a general idea of the type of implications that you would like to draw (guided by a general understanding of the types of conclusion which it is possible to draw within subjectivist theory). You then choose an organizational framework within which such implications may be drawn, and only then do you make the previsions which are required in the given framework. Of course, this is a simplification in that you may find some prevision statements far easier to make than others, your individual specifications may suggest new organizing frameworks, you may have a variety of intentions which change during the specification process and so forth. However, for the most part, you "think" not in individual previsions, as this is too small a unit, but rather in organized collections of previsions (just as you do not "think" in individual words, but rather in sentences, or underlying concepts or some larger unit).

Just as a prevision is a minimal statement of belief, so a belief structure is a minimal organization of beliefs (for purposes of "general reasoning"). Thus, belief structures are objects of fundamental interest. We shall subsequently demonstrate that the arguments that are usually put forward, in imprecise form, as to the content of the theory *can be given* precise expression within this framework. Thus, at the least, this construction is of basic interest in understanding the subjectivist theory. Further, to our knowledge, there is no weaker construction, intermediate between the linear space L and the inner product space A which is rich enough to function as the setting for the theory (rather than simply being useful for an individual problem). Thus, the justification for our choice of A as the basic object of interest will be our repeated demonstration that this construction underlies the basic operation of the theory. However, there is no sense in which you "must" specify your previsions in such a way as to form an inner product space. All that we argue is that a variety of useful and interesting consequences follow quite generally from such a specification. In our opinion, these consequences reveal quite clearly the possibilities and limitations *of the subjectivist theory.*

(6) Belief Construction

To clarify the above ideas, consider a simple scenario which we shall expand in later articles to illustrate our various constructions. In this scenario, a teacher is wondering whether to advise or allow a certain student to register for a particular course. (Say this course is a second year course, and the student has just completed the first year.) The question at issue is whether the student will perform sufficiently well to justify registering for the course, or whether he/she might be better advised to take an alternative "safer" course. Thus, the teacher needs to assess his attitudes to the future performance of the student, and he may have to report his conclusions, (along with his reasons for these conclusions), to a wider community, (say the student, or fellow faculty members). He also expects to encounter similar problems quite frequently in the future, so he is interested in the extent to which his assessment may be reduced to a "routine".

As the teacher considers this problem, perhaps he will list various criteria for success or failure in the course, and a variety of other features, such as reports by fellow teachers, previous test scores, the general background of the student, the difficulty of the course, and so forth. Typically, there will be a bewildering variety of sources of uncertainty. The subjectivist theory cannot remove the difficulty by providing automatic routines to handle the problem. The teacher will have to make his own assessments based on his judgements and observations. All that theory provides is an organizing framework for aspects of thought - it does not pretend to provide the thinking.

Suppose that there will be a test at the end of the year, and the test score S will serve as the final assessment. Thus, as S is the basic quantity of interest, he sets the first element, X_1 in $\mathcal{C}$, to be $X_1 = S$. To assess variability of his assessment, he sets $X_2 = S^2$. Because there is a particular score s below which the student will automatically fail, he sets $X_3 = I(S > s)$, (i.e. $X_3 = 1$ if $S > s$, otherwise $X_3 = 0$). He will be influenced in this judgement by the student's first year test scores. He sets $X_4 = Y$, where Y is the test score on a first year course which is somewhat similar to the proposed course and he set $X_5 = Y^2$. He sets $X_6 = Z$, $X_7 = Z^2$ where Z is the average of the tests scores on the remaining courses. To assess interactions between specific and general ability, he sets $X_8 = YZ$. Finally, he also receives a report from a first year tutor, which includes a numerical rating of ability W, so he sets $X_9 = W$. (At present, the Y, Z and W values are not known by the teacher.)

Note two aspects of the choice of elements of $\mathcal{C}$. Firstly, each element must be a well-defined numerical quantity (which, under some well-defined circumstances, will be revealed). However much we might psychologically prefer to attach previsions to so-called "underlying quantities", such as a student's "innate ability" (perhaps expressed through some model), the prevision must always attach to the result of a measuring process (such as a test). (It is, however, useful and meaningful to construct "models" for "underlying quantities", and we will devote considerable attention in our development to the demonstration that such constructions are possible strictly within the subjectivist theory, i.e. where we may only apply previsions to realizable random quantities.)

Secondly, note that when the teacher has given sufficiently precise meaning to, for example, the first year test score Y, he must then decide which aspects of his belief about Y to specify. He specifies a list $X_{a_1} = f_1(Y), X_{a_2} = f_2(Y), \ldots, X_{a_K} = f_K(Y)$. He may include as many or as few functions as he wishes, and no particular choices have any "special status" over any other particular choices. Similarly, he can choose any functions of any groups of quantities he wishes (e.g. in the above list, he included $X_8 = YZ$). Each function is a different random quantity, and the only sense in which they are related is that two functions of the same random quantity will have certain linear or logical relationships which place constraints on the joint specification of previsions, e.g. in the above list $X_1 = S$, $X_2 = S^2 = X_1^2$, so that

$$P(X_2) = P(X_1^2) \geq P^2(X_1).$$

Let us emphasize that there is no sense in which the teacher can produce a "complete" listing, $\mathcal{C}$, that is a collection so enormous that it must contain all aspects of all quantities of interest. (Remember that each element X_i that he adds to the list involves him in making all of the required prevision statements about that quantity.) He might well decide to produce a more extensive listing than we have given, including more features of the problem, and more aspects (i.e. functions) of the features that he has here included. However, even that list will still be partial, as compared to further specifications that he might make. Just as a prevision statement reveals an aspect of belief, a belief structure also reveals an aspect of belief (although a more elaborate aspect), and must be analyzed as such (rather than by pretending that it is a complete description of belief). The arguments that we shall provide in subsequent articles apply equally to the above listing, any smaller sublisting, or any far more extensive listing. Of course, the more extensive the listing, the more information that can be extracted, but the logical status of the analysis is the same. Thus, it is simply a subjective assessment on your part (outside, but guided by, theory), as to which belief structure will be appropriate for whatever external purpose you have in mind.

Having constructed the list $\mathcal{C}$, he now adds the unit constant X_0, and specifies the values

$$(X_i, X_j) = P(X_i X_j) \quad i,j = 0,1,\ldots,9.$$

Just as he could choose whatever quantities X_i he felt appropriate, so he may specify his previsions to take any values whatever, provided they obey the constraints of coherence, and they correspond to his actual judgement. Some approaches to specification may appear more sensible than others. The more effort that he expends in studying and understanding the uncertainties of the situation, the more his beliefs are likely to be worth studying. These issues are very important but lie outside the theory (except inasmuch as the theory can help to identify which features of the specifications are likely to be important to the analysis and thus can help focus attention, or theory can suggest, by analogy, models for belief which it may be helpful to explore, and so forth).

Indeed, it is precisely because we take as primitive your ability to make statements of prevision that the theory "works" (i.e. escapes from tautology). Otherwise, we would be forced into the attempt to explain your probability statements by more fundamental "underlying" features of the problem, i.e. we would be forced into the impossible position of claiming that beliefs are explicitly reducible to a set of primitive logical explanations.

Thus, we postpone consideration of the problem of specification until we have specific operations to perform upon the belief structure, in later articles, which will provide informal guidelines as to the important aspects of the specification.

(7) Belief construction and Bayesian specification

The specification requirements of the belief structure are far more modest than the requirements of the usual Bayesian formalism. For example, the entry on Bayesian Inference, by Lindley, in the Encyclopedia of Statistical Sciences (Wiley, 1982), begins as follows.

"According to the Bayesian view, all quantitites are of two kinds: those known to the person making the inference and those unknown to the person; the former are described by their known values, the uncertainty surrounding the latter being described by a joint probability distribution for them all ... Within the Bayesian framework <u>all</u> calculations are performed by the probability calculus, using the probabilities for the unknown quantities."

This quotation is in close accord with the work that is actually done in the Bayesian field. That is, any study entitled "A Bayesian analysis of ..."

will generally carry one, or several joint probability distributions for all quantities of interest (or an apology for failing to provide one). More to the point, there is no alternative essentially weaker level of specification for which theory has been developed.

If we imposed Bayesian specification requirements upon the belief structure, then we would have to proceed as follows. In the illustration of the preceding subsection, the teacher identified four quantities of interest S, W, Y, Z. He would now be required to list all possible combinations $(S = s, W = w, Y = y, Z = z)$, that the quantities could take, and assign a probability to each such outcome. Equivalently, he would be required to include in his list $\mathcal{C}$ of random quantities, every bounded function $f(S,W,Y,Z)$, of the four quantities. Having constructed this list, he would then have to specify his values $P(X_i X_j)$, for all possible choices X_i, X_j of bounded functions.

This is an incredibly daunting task. Indeed, we might suspect that in practice it is impossible. Thus, if we decide to make this a fundamental requirement of the theory, we must carefully explain both why we believe that such a task is possible and also why we believe that it is necessary.

To our knowledge, there has never been a justification of the proposition that you have the ability to make arbitrarily large collections of probability statements, under quite general circumstances. What is usually argued in detail is that you may make any individual assessment of any single probability statement that you require. It is then assumed without argument that to make a very large (even uncountably infinite!) collection of such assessments is "similar", apart from in some purely technical sense.

The effect of treating your limitations as a technical rather than a fundamental problem is to completely remove from the theory any considerations of your actual ability to specify previsions. As a consequence, no-one actually specifies their previsions, or probabilities, except in the most trivial of problems. Instead, there is a small collection of convenient multidimensional probability distributions, which are simple to manipulate and some combination of these is assumed to be your probability specification.

In other words, between the conflicting requirements (i) probability specifications are statements of your belief and (ii) you must specify full probability distribution, the technical requirement (ii) always takes precedence over the fundamental conceptual requirement (i). This is an inevitable process, because a specification requirement that is beyond your ability (or interest) to adhere to actually absolves you from the responsibility to specify any of your beliefs. (After all, what criteria can you use to judge how near you are to an unobtainable goal?)

As this requirement is so damaging to the theory, we must therefore ask as to whether there is some corresponding advantage to be gained from such a specification, which somehow offsets the apparent disadvantages. When we argued that you should arrange your previsions into a belief structure, the justification that we gave was that this would allow you to work at a "higher conceptual level", by providing a variety of fundamentally important operations which would be much more difficult to express without the further structure. The only possible justification for requiring you to expand the list C to the enormous degree of a full probablity specification would be if this would, in some sense, raise you to a yet higher conceptual level on which a further collection of analyses was possible which could not be formulated within our simpler notion of a belief structure.

However, we will see in our subsequent treatment that there is no such additional justification. That is, when we describe the various operations that we shall perform upon the belief structure, the various features of the usual Bayesian analysis will simply appear as special cases, of no different logical status than that for the same analysis applied to any other belief structure.

Thus, as we can offer no possible justification of the requirement that you should specify every possible function of every possible quantity, we will simply analyze the general belief structure, in which you specify those aspects which seem important to you, pointing out, as we go, the ways in which the Bayesian analysis is a special case of this approach. However, as with all aspects of our formulation, this is not a matter for debate. The arguments that we present will be available if you specify a belief structure. If, in any particular instance, you specify a full probability structure, then our arguments will simply apply to the belief structure implied by your specification. It will be helpful to have some notation which makes this link explicit.

Thus, the usual Bayes specification concerns a probability space Ω, and a prior probability measure P over Ω. The implied belief structure A, in this case, is simply the space usually denoted $L^2(\Omega,P)$, the Hilbert space of all square integrable functions over Ω (w.r.t.P), with inner product

$$(f,g) = \int_{\Omega} f(w)g(w)\, dP(w).$$

Because such a belief structure is often quite simple to manipulate, it may be helpful in certain circumstances to use the Bayesian specification for formally modelling your actual belief structure. This may be a useful technique provided that it is viewed as a methodological simplification of, and approximation to, a meaningful subjectivist analysis of beliefs. However, as with any model, it becomes dangerous precisely when the study of the model is thought to be more important than the study of those quantities which are being modelled. In most fields, there is some constraint on such confusion, as the quantities being modelled have a tangible presence which can "falsify" the model. However, the Bayesian "model" has no such obvious possibilities for explicit falsification, so that confusion between the model and the quantities being modelled is, unfortunately, a very common occurence.

Thus, in conclusion, let us emphasize once more that each assertion of a prevision statement is a fact. (That is, it is a fact that you hold the belief corresponding to the prevision statement.) This is what justifies the combination of your prevision statements with other "facts" (such as observational data) to provide further "factual conclusions". It makes no more sense to pretend that you have specified previsions which you have not even considered than it would to pretend that you had seen data which has not even been recorded.

References

de Finetti B. (1974, 1975) Theory of Probability. Vol. 1, 2: London: Wiley.

Goldstein, M. (1981) Revising Previsions: A Geometric Interpretation. J. Roy. Statist. Soc. B, 43, 105-130.

Goldstein, M. (1983) The Prevision of a Prevision. J. Amer. Statist. Assoc., 78, 817-819.

Goldstein, M. (1984a) Separating Beliefs. (To appear in the volume in Bayesian Statistics and Econometrics in Honor of Professor de Finetti.)

Goldstein, M. (1984b) Temporal Coherence. (To appear in the proceeding of the second international conference on Bayesian Statistics at Valencia.)

Goldstein, M. (1984c) Exchangeable Situations. (In submission.)

MG/jk

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 2738	2. GOVT ACCESSION NO.: AD-A147526
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Belief Structures	5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Michael Goldstein	8. CONTRACT OR GRANT NUMBER(s): DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709	12. REPORT DATE: August 1984
	13. NUMBER OF PAGES: 22
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Belief; coherence; prevision; projection

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A belief structure is an inner product space which expresses one's (probabilistic) beliefs about certain random quantities. This structure is a fundamental element of the subjectivist theory of probability. This report provides an essentially non-technical description of the basis of the construction, and the ways in which it differs from more familiar Bayesian constructions.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)