Averaging and Eliciting Expert Opinion*

Peter M. Williams

Department of Computer Science
University of Sussex, Falmer
Brighton, Sussex BN1 9QH

1 August 1987
Revised 10 May 1988

Abstract

The paper considers the problem of averaging expert opinion when opinions are expressed quantitatively by belief functions in the sense of Glenn Shafer. Practical experience shows that experts usually differ in their exact quantitative assessments and some method of averaging is necessary. A natural requirement of consistency demands that the operations of averaging and combination, in the sense of Dempster's rule, should commute. Experience also shows that symmetric belief functions are difficult to distinguish in practice. By forming a quotient of the monoid of belief functions modulo the ideal of symmetric belief functions, we are left with an Abelian group with a natural scalar multiplication making it a real vector space. The elements of this quotient space correspond to what we call "regular" belief functions. This solves the averaging problem with arbitrary weights. The existence of additive inverses for regular belief functions means that contrary evidence can be treated without assuming the existence of complements. Opinions expressed by conditional judgements can be incorporated by lifting suitable measures from a quotient space to a numerator. The appendix describes a computer program for implementing these ideas in practice.

*Preparation of this paper was supported by SERC grant GR/E 05360. The ADRIAN project was sponsored by ICI Pharmaceuticals. Thanks are due to both organisations.

Contents

1 Introduction

Our starting point is the theory of belief functions of Shafer [11]. This is presented in Section 2 as the theory of probability measures on semilattices. The presentation there presupposes no prior knowledge of the theory. The following introductory remarks, on the other hand, are addressed principally to those with a working knowledge of the theory. Others may nonetheless find them helpful if they understand a belief function to be a certain sort of probability measure on a lattice.

1.1 Averaging

The immediate stimulus for the present theory was an attempt to apply the theory of belief functions to the problem of analysing possible adverse drug reactions.[1] A panel of over 40 experts in the field of adverse drug reaction reporting was invited to assess what would be the quantitative impact of various items of evidence on the issue of possible drug involvement, assuming these items of evidence applied in a specific case. Although experts usually agreed on the direction in which each item would point, there was inevitably some disagreement on the extent to which an item of evidence would weigh in favour of or against a given hypothesis. Since the opinion of each expert was represented by a belief function, this led to the problem of finding a suitable way of averaging belief functions.

Belief functions cannot be averaged by combining them in the sense of Dempster's rule. Any average of N copies of the same quantity must be that same quantity whereas Dempster's rule is clearly not idempotent. A candidate for the averaging process which satisfies this requirement is the point-wise arithmetic mean. This always produces a belief function, with or without equal weights. But it fails to satisfy a necessary requirement of consistency.

Suppose that M independent items of evidence are to be assessed for their impact on the hypotheses in some domain. Let the j'th expert's assessment of the impact of the i'th item of evidence be expressed by the belief function $Bel(i, j)$. This gives an array of belief functions. Since the items of evidence are independent, suppose their joint impact to be expressed by Dempster's rule of combination. For each expert $j = 1, \ldots, N$ let $Bel(-, j)$ denote the combination of the $Bel(i, j)$ over items of evidence $i = 1, \ldots, M$. $Bel(-, j)$ represents the j'th experts opinion of the impact of the combined evidence. For each item of evidence $i = 1, \ldots, M$ let $Bel(i, -)$ denote the average of the $Bel(i, j)$ over experts $j = 1, \ldots, N$. $Bel(i, -)$ represents the average over experts of the impact of the i'th item of evidence. We now have two ways of obtaining the final result. We may either average the $Bel(-, j)$ over experts or combine the $Bel(i, -)$ over items of evidence. Consistency demands that

[1] The results are reported in Castle *et al* [3].

we obtain the same result in either case. In other words the operations of averaging and combining should commute.

One reason for insisting on this is that too great a strain is otherwise put on the concepts of the identity of an expert or the identity of an item of evidence. Many individual opinions can be considered to be really an average over time or a simultaneous average over different points of view. Each expert is, as it were, an average of experts. Equally an individual item of evidence can often be considered to be in principle a combination of simpler independent items.

It is not difficult to see that the process of averaging by the arithmetic mean, whether with equal weights or not, fails to commute with combination. Reflection shows that a more promising line of attack is to consider the *geometric* mean of "commonality" functions rather than the *arithmetic* mean of belief functions. The only obstacle is that the set of commonality functions is not closed under the formation of geometric means, even setting aside the renormalisation associated with Dempster's rule. Nonetheless if there is to be a solution, this appears to be the only way forward. If there is no solution it is difficult to see how to apply the theory of belief functions in practice.

1.2 Symmetric Belief Functions

In the theory of additive probability measures over finite Boolean algebras the notions of a symmetric measure and a measure with uniform density coincide. Let X be a non-empty finite set with $|X| = n$ and let $B = PX$ be the power set of X considered as a Boolean algebra. A probability measure on B is invariant under all automorphisms of B if and only if its density is invariant under all permutations of X. This means that the density of a symmetric measure must be constant, or uniform, with the value $1/n$ at each point.

In the theory of belief functions, on the other hand, there are infinitely many symmetric measures over a power set with at least two elements. If $X = \{a, b\}$, for example, any belief function whose basic probability assignment satisfies $m(a) = m(b)$ is symmetric. The question arises whether these various symmetric measures can be distinguished in practice.

We can connect this question with the previous one as follows. Suppose that $X = \{a, b\}$ and that we ask two experts to assess the impact of a given item of evidence on the propositions in $PX = \{\{a, b\}, \{a\}, \{b\}, \{\}\}$. Suppose that one expert considers it to support $\{a\}$ to degree 0.1 and the other considers it to support $\{b\}$ to the same degree 0.1. What is the average of their opinions? It must be some symmetric measure. But which one?

Here is an example. Suppose we are concerned with the question whether or not a drug administered to a patient was the cause of a certain adverse event. One factor amongst many which might be relevant is the frequency with which a physician sees patients with this complaint in any case during the course of general practice. Experts agree that if patients present with these symptoms roughly once a day on average, this counts as evidence against the hypothesis of drug involvement. If patients present with these symptoms only very infrequently, say once every five years, this would count as evidence in favour of drug involvement. But there are cases between these extremes of 'once a day' and 'once every five years'. Consider the case of 'once every six months'. One can imagine two experts disagreeing on their assessment of this frequency. One thinks it on the common side of normal and the other thinks it on the uncommon side of normal. Suppose that one expert thinks it supports the hypothesis of drug involvement to degree 0.1 whilst the other thinks it goes against the drug hypothesis to the same degree of 0.1. How should we represent their average opinion?

Now we can imagine one and the same expert similarly being in two minds as it were. One thing, however, is clear to our expert. On balance the evidence is neutral. Its impact must be expressed by a symmetric belief function. But which one? We can put the point in a practical way by considering the problem of the "knowledge engineer" in eliciting an opinion. What could be said to help the expert choose one out of these infinitely many symmetric belief functions? The suggestion which emerges in this paper is that there is no practical significance to this question. The only clear fact in such a case is that the opinion is symmetric between the two alternatives. There is room, therefore, for essentially only one symmetric belief function with practical significance and it may as well be the vacuous one.

This line of thought will lead us to identify all symmetric belief functions. But this has far reaching implications. It means we must identify any two belief functions which "differ" only by a symmetric belief function. In short we shall form a quotient of the commutative monoid of belief functions (the monoid operation corresponding to Dempster's rule) by factoring out the ideal of symmetric belief functions. Fortunately each equivalence class contains one and only one "regular" measure so that we always have a concrete representative to deal with. The quotient monoid turns out to be an Abelian group with a natural scalar multiplication making it a real vector space. The vector space structure leads immediately to a solution of the averaging problem. The existence of additive inverses allows us to treat the problem of contrary evidence without assuming the existence of complements.

1.3 Contrary Evidence

We commonly speak of an item of evidence as being either favourable or unfavourable to a hypothesis. In the theory of belief functions, unfavourable or contrary evidence is treated as a derivative concept, as it is in the ordinary additive theory. Evidence against a proposition is considered to be equivalent to evidence in favour of its formal negation. There appear to be a number of grounds for questioning this assumption.

In the first place it accords poorly with our intuitive ways of thinking. Here is an example. An important factor in assessing possible adverse drug reactions is the time lapse between administration of a drug and onset of symptoms. If the time to onset is too great, this goes against the hypothesis of drug involvement. Intuitively it does so directly and not by virtue of being evidence in favour of some alternative hypothesis about the cause of the symptoms. To take a more homely example, if your friend considers that her present indigestion may have been caused by some prawns she has eaten, discovery that the prawns were eaten three weeks ago may lead you to count this as evidence against her hypothesis. If the prawns were going to give her an upset stomach, they would have done so long before now. The form of reasoning is clear. If the hypothesis were true, something else would be expected to follow. That consequence has failed to materialise (in the time expected). Hence the evidence goes directly against the hypothesis. It would be unnatural and should be unnecessary to interpolate the formal negation of the hypothesis into the argument.

A second reason for questioning the assumption is more technical. Suppose we have two independent items of evidence relating to a given hypothesis. One weighs in favour of the hypothesis to a certain degree and the other weighs against it to the same degree. These two items of evidence ought to cancel out. They point in opposite directions and to the same extent. At least they should combine to form a symmetric belief function. Yet this never happens if we follow the approach of the theory of belief functions, except in the single case of a two-element power set.

In the theory to be presented below, the idea of contrary evidence is based on the intuition of the preceding paragraph, supported by the existence of additive inverses in the quotient monoid of regular belief functions. If we can agree on a regular belief function to express the impact of evidence favourable to a given hypothesis to a certain degree, evidence unfavourable to that hypothesis to the same absolute numerical degree will be represented by the additive inverse of the first belief function. By definition the two will cancel out when combined.

None of this is to deny that we may sometimes have evidence directly in favour of each of two complementary hypotheses or directly against each of two complementary hypotheses. But neither situation is the same as having evidence directly in favour and directly against just one.

2 Probability Measures on Inflattices

In [11] and most later expositions belief functions are defined on power sets. A central tenet of this paper is that the way to view the theory of belief functions is as part of an autonomous theory of probability measures on semilattices, even when they happen to be Boolean algebras, rather than as a competitor to the orthodox theory of additive probability measures. Further discussion of this point will be found in Section 4. We begin with an outline of the necessary lattice-theoretic background. For more details see, for example, [8, Ch.I].

2.1 Distributive Lattices

A *partially ordered* set is a set A with a binary relation $\leq$ satisfying

$$a \leq a$$

$$a \leq b \text{ and } b \leq a \text{ imply } a = b$$

$$a \leq b \text{ and } b \leq c \text{ imply } a \leq c$$

for all a, b, c in A. A subset S of a partially ordered set A is said to be an *upper set* if it is closed above, *i.e.* if $a \in S$ and $a \leq b$ imply that $b \in S$. Dually a *lower set* is any set which is closed below. We write $\uparrow(a) = \{b \in A | a \leq b\}$ for the smallest upper set containing a and $\downarrow(a) = \{b \in A | b \leq a\}$ for the smallest lower set containing a.

A *lattice* is a partially ordered set in which every finite subset has both a greatest lower bound, called its *meet*, and a least upper bound, called its *join*. We write $\bigwedge S$ and $\bigvee S$ for the meet and join of the set S when they exist. When S is a two element set we write $a \wedge b$ instead of $\bigwedge\{a, b\}$ and $a \vee b$ instead of $\bigvee\{a, b\}$. Thus $a \vee b$ and $a \wedge b$ are determined when they exist by the conditions

$$a \vee b \leq c \text{ iff } a \leq c \text{ and } b \leq c$$

$$c \leq a \wedge b \text{ iff } c \leq a \text{ and } c \leq b$$

for all $c \in A$. Since the empty set is certainly finite it follows that a lattice has both top and bottom elements providing a meet and a join, respectively, for the empty set. We refer to these as 1 and 0. Equivalently, a lattice is a partially ordered set with binary meets and binary joins and a top and bottom element.

An element a of a lattice L is *complemented* when there is an element b of L such that

$$a \wedge b = 0 \text{ and } a \vee b = 1.$$

In that case b is said to be a complement of a. In general an element of a lattice may have many complements. For an important class of lattices, however, a given element can have at most one.

A lattice L is *distributive* if it satisfies either of the equivalent conditions

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c) \qquad (\forall a, b, c \in L)$$

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) \qquad (\forall a, b, c \in L).$$

Complements in a distributive lattice are unique whenever they exist.

A *Boolean algebra* is a distributive lattice in which every element has a complement. Since complements are defined through meets and joins, a morphism of distributive lattices which happen to be Boolean algebras is necessarily a Boolean homomorphism.

2.2 Probability Measures on Distributive Lattices

A probability measure on a distributive lattice D is any real unit-interval valued function $p\colon D \to [0,1]$ satisfying

(1) $p(a \vee b) + p(a \wedge b) = p(a) + p(b)$

(2) $a \leq b$ implies $p(a) \leq p(b)$

(3) $p(0) = 0$ and $p(1) = 1$.

When D is a Boolean algebra this agrees with the usual definition of a probability measure on a Boolean algebra. More generally we have the following:

Proposition 1 *Every probability measure on a distributive lattice D has a unique extension to a probability measure on the Boolean algebra freely generated by D.*

The meaning of this is as follows. The Boolean algebra freely generated by a distributive lattice D is a Boolean algebra BD together with a morphism $\eta\colon D \to BD$ of distributive lattices which is universal amongst morphisms with this property.[2] Thus if B is any Boolean algebra and $f\colon D \to B$ is a morphism of distributive lattices, there exists a unique Boolean homomorphism $\overline{f}\colon BD \to B$ such that the diagram

$$\begin{tikzcd} D \arrow[r, "\eta"] \arrow[dr, "f"'] & BD \arrow[d, dashed, "\overline{f}"] \\ & B \end{tikzcd}$$

commutes. To say that a probability measure $\overline{p}$ on BD is an extension of a probability measure p on D is to say that $\overline{p}$ restricts to p along the canonical insertion η. We can interpret this as confirming the correctness of the definition of a probability measure on a distributive lattice, relative to the concept of a probability measure on a Boolean algebra.

As is well known (1) can be extended to the meet or join of several elements. Thus reformulating (1) as

$$p(a \vee b) = p(a) + p(b) - p(a \wedge b)$$

[2]If D is already presented as a sub-distributive lattice of a Boolean algebra B, the power set algebra $B = PX$ for example, then BD is just the sub-Boolean algebra of B generated by D. The construction of BD in general is more intricate: see for example [8, II.4.5] or [9, page 25].

leads to

$$p(a \vee b \vee c) = \\ p(a) + p(b) + p(c) - p(a \wedge b) - p(a \wedge c) - p(b \wedge c) + p(a \wedge b \wedge c).$$

In general, if S is any finite subset of a distributive lattice D and p is a probability measure on D, then

$$p(\bigvee S) = 1 - \sum_{R \subseteq S} (-1)^{|R|} p(\bigwedge R)$$

where $|R|$ is the number of elements in R and summation over the empty set is included. This means that the probability of the join of a set of elements is uniquely determined by the probabilities of the meets of all its subsets. Distributivity is essential for deriving this result.

2.3 Semilattices

Lattices are required to have both finite meets and finite joins. If we only require finite limits of one sort, we are dealing with a semilattice. Clearly there are two sorts of semilattice from the order-theoretic point of view. A *meet semilattice* is a partially ordered set in which every finite subset has a meet. A *join semilattice* is a partially ordered set in which every finite subset has a join. A meet semilattice necessarily has a distinguished top element, 1 say, while a join semilattice has a distinguished bottom element 0.

A morphism of meet semilattices is a map which preserves finite meets while a morphism of join semilattices is a map which preserves finite joins. Such maps necessarily preserve order. Notice that morphisms of semilattices must preserve the empty meet or the empty join. Thus a morphism of meet semilattices must preserve the top element but need not preserve the bottom element while a morphism of join semilattices must preserve the bottom element but need not preserve the top element. In general a meet semilattice may well have finite or even arbitrary joins. (It always will if the semilattice is finite or, more generally, complete.) But these need not be preserved by morphisms of meet semilattices. A corresponding remark holds for join semilattices.

A semilattice can also be considered to be a special sort of commutative monoid. The operations of binary join and taking the bottom element of a join semilattice, for example, satisfy the following equations:

(4) $a \vee (b \vee c) = (a \vee b) \vee c$

(5) $a \vee b = b \vee a$

(6) $a \vee 0 = a$

(7) $a \vee a = a.$

The first three define a commutative monoid—a commutative semigroup (4, 5) with unit (6)—in which every element is idempotent (7). The order relation can be recovered by the definition

$$a \leq b \text{ iff } a \vee b = b.$$

A meet semilattice satisfies the same equations if $\wedge$ replaces $\vee$ and 1 replaces 0. In that case the order relation must be recovered in the opposite sense:

$$a \leq b \text{ iff } a \wedge b = a.$$

From the algebraic point of view there is no difference between a meet semilattice and a join semilattice. It is a matter of terminology corresponding to the distinction between an additive and a multiplicative group. The distinction only arises when we choose one of the two senses for the ordering.

A meet semilattice is said to be *complete* if every subset has a meet. Every finite meet semilattice is complete. Dually a join semilattice is complete if every subset has a join. The categories of complete join and complete meet semilattices are particularly rich in structure. Following Joyal and Tierney [9] we call them the categories of *suplattices* and *inflattices* respectively. Their richness arises principally from the fact that each is isomorphic to its opposite as follows.

Let A be a suplattice (complete join semilattice). Then A necessarily has arbitrary meets as well as arbitrary joins. The meet of any subset S is just the join of its lower bounds:

$$\bigwedge S = \bigvee\{a \in A | S \subseteq \uparrow(a)\}.$$

Dually an inflattice has arbitrary joins.[3] This means that whenever A is a suplattice—a partially ordered set with arbitrary joins—the same set with the opposite partial order, A^{o} say, is also a suplattice. The joins of A^{o} are just the meets of A. Dually the opposite of an inflattice is an inflattice. Now let A and B be suplattices and let $f\colon A \to B$ be a morphism of suplattices, so that f preserves arbitrary joins. Then there is a unique morphism of partially ordered sets $f_*\colon B \to A$ satisfying the condition

$$f(a) \leq b \text{ iff } a \leq f_*(b)$$

for all $a \in A$ and $b \in B$. Explicitly f_* is given by

$$f_*(b) = \bigvee\{a \in A | f(a) \leq b\}.$$

f_* preserves arbitrary meets since A has arbitrary joins and f preserves them. In general if $f\colon A \to B$ and $g\colon B \to A$ are order-preserving maps between partially ordered sets and

$$f(a) \leq b \text{ iff } a \leq g(b)$$

[3]Thus, as an object, a complete semilattice or either sort is in fact a complete lattice. However, since a morphism of suplattices need not preserve meets, nor a morphism of inflattices joins, the categories of complete semilattices and complete lattices have significantly different categorical properties.

for all $a \in A$ and $b \in B$, we refer to f as the *left adjoint* of g and g as the *right adjoint* of f. The left adjoint preserves joins and the right adjoint preserves meets. Either uniquely determines the other by the condition

$$f(a) = \bigwedge\{b \in B | a \leq g(b)\}$$

$$g(b) = \bigvee\{a \in A | f(a) \leq b\}.$$

Now meets in A are precisely the joins of A^{o} so that f_* can equally be considered as a morphism of suplattices $f^{\mathrm{o}}: B^{\mathrm{o}} \to A^{\mathrm{o}}$. This establishes a bijection between the suplattice morphisms from A to B and the suplattice morphisms from B^{o} to A^{o}. In fact the functor which sends a suplattice to its opposite and a morphism to its right adjoint establishes an isomorphism between the category of suplattices and its opposite. Dually the functor which sends an inflattice to its opposite and a morphism to its left adjoint establishes an isomorphism between the category of inflattices and its opposite.

From an algebraic point of view our concern in this paper is with probability measures on complete semilattices in general. Our immediate practical concern, however, is with probability measures on finite semilattices. To avoid complications, we shall therefore restrict attention to the full subcategories of finite suplattices and finite inflattices.

2.4 Probability Measures on Inflattices

Definition 1 *A* probability measure *on a finite inflattice A is a real unit-interval valued function $p: A \to [0,1]$ satisfying*

$$p(\bigvee S) + \sum_{R \subseteq S} (-1)^{|R|} p(\bigwedge R) \geq 1$$

for every (finite) subset $S \subseteq A$.

Lemma 2 *Let $f: A \to B$ be a morphism of finite inflattices and let q be a probability measure on B. Define $p: A \to [0,1]$ by*

$$p(a) = q(f(a))$$

for all $a \in A$. Then p is a probability measure on A, which we denote by the functional composition $q \circ f$.

This follows from the definition and a relatively simple though not trivial combinatorial argument. Our aim now is to show that the definition of a probability measure on a meet semilattice is justified by the relation between probability measures on meet semilattices and probability measures on the distributive lattices which they freely generate.

The distributive lattice freely generated by a finite inflattice A is the lattice DA of all lower sets of A ordered by inclusion. The insertion of generators is the down-segment map $\downarrow: A \to DA$ which sends an element

a of A to $\downarrow(a)$. DA is a distributive lattice since it is a sublattice of the power set lattice of A which is evidently distributive. The down-segment map preserves meets since

$$a \leq b \wedge c \text{ iff } a \leq b \text{ and } a \leq c$$

and $\downarrow(1) = A$. Moreover it is universal with this property. If B is any (finite) distributive lattice and $f\colon A \to B$ preserves finite meets, there is a unique morphism of distributive lattices $\overline{f}\colon DA \to B$ such that $f = \overline{f} \circ \downarrow$.

Now let $\overline{p}$ be a probability measure on DA for some finite inflattice A. Evidently $\overline{p}$ is a probability measure on DA considered as an inflattice. Since $\downarrow\colon A \to DA$ is a morphism of inflattices the function p defined by $p(a) = \overline{p}(\downarrow(a))$ is a probability measure on A by Lemma 2. We say then that p is the restriction of $\overline{p}$ and $\overline{p}$ is an extension of p along the map $\downarrow$. The fundamental result on probability measures on inflattices says that every probability measure can be had uniquely in this way.

Proposition 3 *Every probability measure on a finite inflattice A has a unique extension to a probability measure on the distributive lattice freely generated by A.*

The proposition contains two parts: the existence of the extension and its uniqueness. Uniqueness is straightforward. If $\overline{p}$ is an extension of p then $\overline{p}$ is already determined on all principal lower sets of the form $\downarrow(a)$ as a ranges over A. Every element of DA is a finite union of such sets. But we already know that the probability of the join (union) of a finite set of elements of a distributive lattice is uniquely determined by the probabilities of the meets (intersections) of all its subsets, and if R is any finite subset of the meet semilattice A, then

$$\bigcap_{a \in R} \downarrow(a) = \downarrow(\bigwedge R).$$

Thus the extension, if it exists, must be given uniquely by

$$\overline{p}(\bigcup_{a \in S} \downarrow(a)) = 1 - \sum_{R \subseteq S} (-1)^{|R|} p(\downarrow(\bigwedge R))$$

for every finite subset $S \subseteq A$. The proof of existence is less immediate. It must be verified that $\overline{p}$ is well-defined and that it is a probability measure on the distributive lattice DA. This follows from combinatorial results of Revuz [10] inspired by Choquet [4]. (See also [7].)

The preceding proposition allows us to introduce the idea of a dual probability measure on the opposite inflattice. This corresponds to the "commonality function" of the theory of belief functions.

Proposition 4 *Suppose p is a probability measure on a finite inflattice A and that $\overline{p}$ is its unique extension to the distributive lattice DA. Define p° for all a in A by*

$$p^{\circ}(a) = 1 - \overline{p}(\uparrow(a)^{c})$$

where $\uparrow(a)^{\mathrm{c}}$ *is the set-theoretic complement of* $\uparrow(a)$ *as a subset of* A*. Then* p^{o} *is a probability measure on the inflattice* A^{o} *and moreover the correspondence* $p \mapsto p^{\mathrm{o}}$ *is bijective.*

Proof For every upper set S of A define

$$q(S) = 1 - \overline{p}(S^{\mathrm{c}}).$$

Since the complement of an upper set is a lower set, q is a well-defined function on the collection UA of upper sets of A. The latter form a distributive lattice and it is easy to verify that q is a probability measure on UA and that p^{o} is the restriction of q to A along the map $\uparrow: A \to UA$. But this map preserves the joins of A and hence the meets of A^{o}. Thus p^{o} is a probability measure on the inflattice A^{o} by Lemma 2. To establish bijectivity, note that the upper sets of A are precisely the lower sets of A^{o}, in other words $UA = D(A^{\mathrm{o}})$. Hence p^{o} has a unique extension $\overline{p^{\mathrm{o}}}$ to a probability measure on $D(A^{\mathrm{o}})$. But q is an extension of p^{o} to a probability measure on the distributive lattice $D(A^{\mathrm{o}})$, since p^{o} is by definition the restriction of q. So q must be this unique extension, in other words $\overline{p^{\mathrm{o}}} = q$, from which it follows[4] that $p^{\mathrm{oo}} = p$. □

Proposition 3 provides a useful representation of probability measures on finite semilattices. We know that every probability measure on an inflattice A extends uniquely to a probability measure on the distributive lattice DA of lower sets of A. This in turn extends uniquely to a probability measure on the Boolean algebra BDA freely generated by DA. But when A is finite, the Boolean algebra BDA freely generated by DA is just the the power set algebra PA of all subsets of A. This is because DA is a sub-distributive lattice of the Boolean algebra PA and every upper set of A occurs in BDA as the complement of a lower set and hence every singleton $\{a\}$ occurs in BDA as the intersection of the upper set $\uparrow(a)$ and the lower set $\downarrow(a)$. Finally every subset of A occurs in BDA as a (finite) union of singletons. This means that, for every probability measure p on a finite inflattice, there is a unique probability measure on the Boolean algebra PA which restricts via the down-segment mapping to p. But every probability measure on a finite power set algebra is uniquely determined by its values on singletons, or in other words by its density. Putting these observations together we have the following:

Proposition 5 *A real-valued function* p *on a finite inflattice* A *is a probability measure on* A *if and only if there exists a real unit-interval valued function* $m: A \to [0,1]$ *with* $\sum_{a\in A} m(a) = 1$ *such that*

$$p(a) = \sum\{m(b) | b \leq a\}$$

[4]In the general case this correspondence between probability measures on an inflattice and probability measures on its dual is bijective only if probability measures satisfy an additional continuity condition. Compare the condition of "condensability" of [13]. In that case Proposition 3 has to concern itself with extensions to probability measures on the *frame* freely generated by the inflattice: see [8, pages 39–41].

for all $a \in A$. Moreover this function, called the density *of p, is unique when it exists.*

It is easy to see that a probability measure p and its dual p^{o} have the same density. Thus if m is the density of p then p^{o} is given by

$$p^{\mathrm{o}}(a) = \sum\{m(b) | a \leq b\}.$$

Using the representation in terms of densities it is easy in the finite case to establish an important result concerning measures on direct sums of inflattices. Let A and B be inflattices. Denote by $A \oplus B$ the set of all pairs (a, b) with $a \in A$ and $b \in B$. Then $A \oplus B$ is an inflattice under the coordinatewise partial order. It is in fact the biproduct of A and B in the category of inflattices under the obvious injections and projections.

Proposition 6 *If p and q are probability measures on the finite inflattices A and B respectively, then the function $p \times q$ defined for all $a \in A$ and $b \in B$ by*

$$(p \times q)(a, b) = p(a)q(b)$$

is a probability measure on $A \oplus B$.

Proof The density m of $p \times q$ is just the pointwise real product of the densities of p and q. Thus $m(a, b) = m_p(a)m_q(b)$ where m_p is the density of p and m_q is the density of q. $\square$

Corollary 7 *If p and q are probability measures on an inflattice A then the function $p \cdot q$ defined for all $a \in A$ by*

$$(p \cdot q)(a) = p(a)q(a)$$

is also a probability measure on A.

Proof Let $\Delta: A \to A \oplus A$ be the diagonal morphism sending a to (a, a). Then $p \cdot q = (p \times q) \circ \Delta$ which is a probability measure on A by Lemma 2. $\square$

Corollary 7 allows us to introduce a binary operation on probability measures on an inflattice which forms the basis of Dempster's rule of combination in the theory of belief functions. Let $Pr(A)$ denote the set of probability measures on an inflattice A and let the binary operation $\star$ be defined by

$$p \star q = (p^{\mathrm{o}} \cdot q^{\mathrm{o}})^{\mathrm{o}}$$

for all $p, q \in Pr(A)$.

Proposition 8 *$Pr(A)$ is a commutative monoid under $\star$.*

Proof Associativity follows from associativity of real multiplication and the fact that $p^{oo} = p$ and commutativity is obvious. The unit of the monoid is the measure whose dual has the constant value 1. □

The unit v of $Pr(A)$, which we call the *vacuous* measure on A, is given explicitly by

$$v(a) = \begin{cases} 1 & \text{if } a = 1 \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to show that if p has density m_p and q has density m_q then $p \star q$ has density[5]

$$m(a) = \sum\{m_p(b)m_q(c) | a = b \wedge c\}.$$

Now let $f\colon A \to B$ be an inflattice morphism and let f^o denote its left adjoint as a morphism of the opposites B^o and A^o considered as inflattices. Then we define $Pr(f)\colon Pr(A) \to Pr(B)$ by

$$Pr(f)(p) = (p^o \circ f^o)^o$$

for all $p \in Pr(A)$.

Proposition 9 *Pr is a (covariant) functor from the category of finite inflattices to the category of commutative monoids.*

Proof If $p \in Pr(A)$ and $f\colon A \to B$ is an inflattice morphism, it follows that $f^o\colon B^o \to A^o$ is also an inflattice morphism and hence by Lemma 2 that $p^o \circ f^o \in Pr(B^o)$ since $p^o \in Pr(A^o)$. Thus $(p^o \circ f^o)^o = Pr(f)(p) \in Pr(B)$. Now suppose that $p, q \in Pr(A)$. Then

$$\begin{aligned} Pr(f)(p \star q) &= ((p \star q)^o \circ f^o)^o \\ &= ((p^o \cdot q^o) \circ f^o)^o \\ &= ((p^o \circ f^o) \cdot (q^o \circ f^o))^o \\ &= (p^o \circ f^o)^o \star (q^o \circ f^o)^o \\ &= Pr(f)(p) \star Pr(f)(q) \end{aligned}$$

and $Pr(f)$ evidently preserves the unit of $Pr(A)$. Thus $Pr(f)$ is a monoid homomorphism. To see that $Pr(g \circ f) = Pr(g) \circ Pr(f)$ for any inflattice morphisms $f\colon A \to B$ and $g\colon B \to C$ we only need observe that $(g \circ f)^o = f^o \circ g^o$ and Pr evidently preserves identities. □

The probability functor Pr has a simple description in terms of densities. If p is a probability measure on A with density m and $f\colon A \to B$ is a morphism of inflattices then the density m_f of $Pr(f)(p)$ is given for all $b \in B$ by

$$m_f(b) = \sum\{m(a) | f(a) = b\}.$$

[5]This means that $\star$ is the restriction to the densities on A of the multiplication in the Möbius algebra of A. See, for example, [15] and [5].

To see this, let p_f be the measure on B with density m_f. Then

$$\begin{aligned}(p_f)^{\mathrm{o}}(b) &= \sum\{m_f(c)|b \leq c\}\\ &= \sum\{m(a)|b \leq f(a)\}\\ &= \sum\{m(a)|f^{\mathrm{o}}(b) \leq a\}\\ &= p^{\mathrm{o}}(f^{\mathrm{o}}(b)).\end{aligned}$$

Hence $(p_f)^{\mathrm{o}} = p^{\mathrm{o}} \circ f^{\mathrm{o}}$ and therefore $p_f = Pr(f)(p)$ from which the result follows by the bijective correspondence between measures and their densities.

3 Regular Measures on Inflattices

Having defined probability measures in general on finite inflattices, our task in line with the thoughts of the introduction is to determine the uniform measures on an inflattice and to examine the structure of the quotient monoid obtained by factoring them out. This turns out to be an Abelian group, each element of which contains a distinguished representative which we call a "regular" measure.

3.1 Uniform Measures

We first introduce a restriction on the monoid of probability measures on an inflattice. We say that a probability measure p on an inflattice A is *proper* if and only if it assigns the maximum value of 1 to the top element alone: thus if

$$p(a) = 1 \text{ implies } a = 1.$$

This is equivalent to saying that p^{o} is strictly positive, from which it follows that the proper probability measures on an inflattice form a submonoid. Correspondingly if $f\colon A \to B$ is any inflattice morphism then the image under $Pr(f)$ of a proper probability measure is proper. Accordingly from now on $Pr(A)$ will denote just the proper probability measures on A and $Pr(f)$ will denote the corresponding restriction. The significance of this restriction is that the monoid $Pr(A)$ of proper probability measures on A now satisfies the cancellation law:

$$p \star q = p \star r \text{ implies } q = r.$$

We say that two elements a and b of a finite inflattice A *have the same rank* if and only if they have the same number of elements above them: thus if $|{\uparrow}(a)| = |{\uparrow}(b)|$. It is clear that this relation partitions the underlying set of A and that the equivalence classes are linearly ordered by the cardinal numbers associated with them, *i.e.* by the number of elements in ${\uparrow}(a)$ for any a in a given equivalence class. Then the equivalence class of lowest rank is just the singleton consisting of the top element alone. The class of highest rank contains just the bottom element. Since the exact number of elements in $\uparrow(a)$ is not important we shall count the ranks from top to bottom by the integers $0, 1, \ldots$ and call this numerical assignment the *rank function*. For a power set inflattice $(PX)^{\mathrm{o}}$ ordered by opposite inclusion and with S a subset of X where $|X| = n$, $\mathrm{rank}(S)$ is just the cardinality of S with $\mathrm{rank}(X) = n$ and $\mathrm{rank}(\emptyset) = 0$. One can think of the rank function as a measure for inflattices of the depth of an element below the top. Since the top element is the unit of an inflattice considered as a monoid, this is the same as a measure of the distance from the unit.

Definition 2 *We define a proper probability measure p on A to be* uniform *if its dual measure p^{o} is constant across ranks. Thus p is uniform iff*

$$\mathrm{rank}(a) = \mathrm{rank}(b) \ \textit{implies} \ p^{\mathrm{o}}(a) = p^{\mathrm{o}}(b).$$

Evidently a probability measure on a free semilattice (a power set or its opposite) is uniform if and only if it is symmetric, *i.e.* invariant under all order-preserving automorphisms. We denote by $Un(A)$ the uniform probability measures on an inflattice A. The unit of the monoid $Pr(A)$ is always uniform since its dual measure has the constant value 1 everywhere. Another uniform measure on a finite inflattice A is the measure whose density is given for all $a \in A$ by $m(a) = 1/n$ where $n = |A|$.

For the next lemma we recall that a subset I of a commutative monoid $M = (M, \star, 0)$ is an *ideal* of M if and only if

(1) I is a sub-monoid of M

(2) $a \in I$ and $a \star b \in I$ imply $b \in I$.

Lemma 10 *$Un(A)$ is an ideal of $Pr(A)$.*

The proof is straightforward. The restriction to proper measures is needed to verify (2).

A subset of a commutative monoid is the kernel of a monoid epimorphism iff it is an ideal. Accordingly we can form the quotient monoid $Pr(A)/Un(A)$ and the kernel of the canonical quotient map is then exactly the ideal $Un(A)$ of uniform measures. Explicitly we say that p is *equivalent* to q, written $p \equiv q$, iff

$$p \star u = q \star v \ \text{ for some } \ u, v \in Un(A).$$

This is a congruence relation since $Un(A)$ is a submonoid and the canonical projection sending each proper probability measure p to its congruence class $[p]$ is a monoid homomorphism. From condition (2) it follows that the unit of the quotient monoid, which is the congruence class determined by the unit of $Pr(A)$, is exactly the set $Un(A)$ of uniform measures. Thus $[0] = Un(A)$.

Our next task is to show that the quotient monoid $Pr(A)/Un(A)$ is in fact a group. Using the symbol $+$ to denote the monoid operation in the quotient, this means that for each measure $p \in Pr(A)$ there is a measure $q \in Pr(A)$ such that $[p] + [q] = [0]$ or in other words such that $p \star q$ is uniform. Since $(p \star q)^{\mathrm{o}} = p^{\mathrm{o}} \cdot q^{\mathrm{o}}$ this means that values of q^{o} across ranks must be proportional to the reciprocals of the values of p^{o}. Thus we need to be able to construct a probability measure with any preassigned ratios of p^{o}-values across ranks. The following result constitutes the basic lemma of the present theory.

Lemma 11 *Let f be any real-valued function on a finite inflattice A with $n+1$ ranks. Then there exists a proper probability measure p on A and a sequence of positive real number $K_0, \ldots, K_n$ such that for each $i = 0, \ldots, n$*

$$p^{\mathrm{o}}(a) = K_i \exp f(a)$$

whenever $\mathrm{rank}(a) = i$.

Proof Assume first that $f(0) = 0$. Define a sequence of functions $(m_i)_{i \leq n}$ where each function m_i is defined on $\{a \in A | \mathrm{rank}(a) \leq i\}$ as follows:

$$m_0(1) = \exp f(1)$$

$$m_i(a) = \begin{cases} k_i m_{i-1}(a) & \text{if } \mathrm{rank}(a) < i \\ \exp f(a) - k_i g_i(a) & \text{if } \mathrm{rank}(a) = i \end{cases}$$

where

$$g_i(a) = \sum \{m_{i-1}(b) | a < b\}$$

$$k_i = \inf \frac{\exp f(a)}{g_i(a)}$$

with the infimum taken over $\{a \in A | \mathrm{rank}(a) = i\}$. It follows by induction that each constant k_i is finite and strictly positive and that each function m_i is non-negative. In particular m_n is non-negative. Let $m = m_n$. Then from the definition we have

$$(1) \quad \sum \{m_i(b) | a \leq b\} = \exp f(a)$$

whenever $\mathrm{rank}(a) = i$, since $a < b$ and $\mathrm{rank}(a) = i$ imply $\mathrm{rank}(b) < i$. Thus $\sum_{a \in A} m(a) = \sum_{0 \leq b} m_n(b) = \exp f(0) = 1$ and so m is a density. Now let

$$K_i = \prod_{i < j \leq n} k_j$$

with $K_n = 1$ so that from the definition

$$(2) \quad m(a) = K_i m_i(a)$$

whenever $\mathrm{rank}(a) = i$. In particular

$$m(1) = K_0 m_0(1) = k_1 \ldots k_n \exp f(1) > 0$$

so that the measure p based on the density m is a proper probability measure. Moreover it follows from (1) and (2) that whenever $\mathrm{rank}(a) = i$ we have

$$p^{\mathrm{o}}(a) = \sum\{m(b)|a \leq b\} = \sum\{K_i m_i(b)|a \leq b\} = K_i \exp f(a)$$

which was the original assertion. If $f(0) \neq 0$, apply the above construction to the function $f'(a) = f(a) - f(0)$ and absorb $\exp -f(0)$ into each of the factors K_i. $\square$

Definition 3 *If f is any real-valued function on a finite inflattice A we denote by* $\mathrm{reg}\, f$ *the proper probability measure defined by the above construction.*

Proposition 12 *$Pr(A)/\,Un(A)$ is an Abelian group.*

Proof Suppose p belongs to $Pr(A)$. Let $q = \mathrm{reg}\,(-\log p^{\mathrm{o}})$. Then by Lemma 11 we have q^{o} proportional to $\exp(-\log p^{\mathrm{o}}) = 1/p^{\mathrm{o}}$ across ranks. Hence $(p \star q)^{\mathrm{o}} = p^{\mathrm{o}} \cdot q^{\mathrm{o}}$ is constant across ranks and therefore $p \star q$ is uniform. Thus $[p] + [q] = [p \star q] = [0]$. $\square$

Our aim now is to show that $Pr(A)/\,Un(A)$ is isomorphic to the additive group of a vector space quotient of the real vector space of all real-valued functions on A. This will allow us to equip $Pr(A)/\,Un(A)$ with a scalar multiplication making it into a real vector space.

If A is a finite inflattice let $L(A)$ denote the free real vector space on the underlying set of A, *i.e.* the space of all real-valued functions on A. Let $N(A)$ denote the subset of $L(A)$ of all functions that are constant across ranks. Thus $f \in N(A)$ iff

$$\mathrm{rank}(a) = \mathrm{rank}(b) \;\text{ implies }\; f(a) = f(b)$$

for all $a, b \in A$. It is clear that $N(A)$ is a vector subspace of $L(A)$. If $f, g \in L(A)$ we write $f \simeq g$ to mean that $f - g$ belongs to $N(A)$, we denote by $[f]$ the equivalence class determined by f, and by $L(A)/N(A)$ the quotient vector space of all such equivalence classes. Our aim is to show that $Pr(A)/\,Un(A)$ is isomorphic to the additive group of $L(A)/N(A)$. We need first the following results:

Lemma 13

(1) *If $p \equiv q$ then* $\log p^{\mathrm{o}} \simeq \log q^{\mathrm{o}}$

(2) *If $f \simeq g$ then* $\mathrm{reg}\, f = \mathrm{reg}\, g$

(3) $p \in Un(A)$ *iff* $\log p^{\mathrm{o}} \in N(A)$

(4) $f \simeq \log\,(\mathrm{reg}\, f)^{\mathrm{o}}$

Proof If $p \equiv q$ there are uniform measures u and v such that $p \star u = q \star v$. Hence $\log p^{\mathrm{o}} - \log q^{\mathrm{o}} = \log v^{\mathrm{o}} - \log u^{\mathrm{o}}$, which belongs to $N(A)$, and (1) follows. The proof of (2) is by induction on rank following the steps of the definition of the function reg in Lemma 11. We leave the details to the interested reader. (3) is a restatement of the definition of uniformity and (4) is a restatement of Lemma 11. □

Proposition 14

$Pr(A)/Un(A)$ is isomorphic to the additive group of $L(A)/N(A)$.

Proof Define $\phi\colon Pr(A)/Un(A) \to L(A)/N(A)$ by

$$\phi([p]) = [\log p^{\mathrm{o}}]$$

and $\psi\colon L(A)/N(A) \to Pr(A)/Un(A)$ by

$$\psi([f]) = [\mathrm{reg}\, f].$$

These are well-defined in view of (1) and (2) above. Now (3) implies that $\phi([p]) = [0]$ iff $[p] = [0]$ and

$$\begin{aligned}
\phi([p] + [q]) &= \phi([p \star q]) \\
&= [\log (p \star q)^{\mathrm{o}}] \\
&= [\log p^{\mathrm{o}} + \log q^{\mathrm{o}}] \\
&= [\log p^{\mathrm{o}}] + [\log q^{\mathrm{o}}] \\
&= \phi([p]) + \phi([q]).
\end{aligned}$$

Thus ϕ is a group homomorphism and moreover injective as a function. Now suppose that f belongs to $L(A)$. Then

$$\phi(\psi([f]) = \phi([\mathrm{reg}\, f]) = [\log (\mathrm{reg}\, f)^{\mathrm{o}}] = [f]$$

by (4). Hence $\phi \circ \psi$ is the identity on $L(A)/N(A)$. Since ϕ is injective it follows that $\psi \circ \phi$ is the identity on $Pr(A)/Un(A)$. Since ϕ is a group homomorphism it follows that ψ is also a group homomorphism and we are through. □

3.2 Regular Measures

For practical purposes it is inconvenient to deal with equivalence classes of measures. We show next that each equivalence class contains a distinguished element which will serve as a canonical representative.

Definition 4 *Let $\rho\colon Pr(A) \to Pr(A)$ be defined by*

$$\rho(p) = \mathrm{reg}\, (\log p^{\mathrm{o}}).$$

We say that a proper measure p is regular *if and only if $\rho(p) = p$ and we denote by $Reg(A)$ the set of regular measures on a finite inflattice A.*

We now show that each element of $Pr(A)/Un(A)$ contains exactly one regular measure.

Lemma 15 *ρ is idempotent: $\rho \circ \rho = \rho$. Hence $\rho(p)$ is regular for all $p \in Pr(A)$.*

Proof Substitute $f = \log p^{\mathrm{o}}$ in (4) above and apply (2). $\square$

Proposition 16 *Each element of $Pr(A)/Un(A)$ contains one and only one regular measure.*

Proof Suppose $[p]$ belongs to $Pr(A)/Un(A)$. Then $\phi(\psi([p])) = [p]$ implies that $[\mathrm{reg}\,(\log p^{\mathrm{o}})] = [p]$ or in other words $[\rho(p)] = [p]$. But $\rho(p)$ is regular by Lemma 15. Hence every element of $Pr(A)/Un(A)$ contains at least one regular measure. Now suppose that $p \equiv q$. Then $\rho(p) = \rho(q)$ by (1) and (2) of Lemma 13. So if p and q are both regular we have $p = \rho(p) = \rho(q) = q$. Thus each equivalence class in $Pr(A)/Un(A)$ contains at most one regular measure. $\square$

This result means that $Reg(A)$ is likewise an Abelian group under the obvious definition of addition. Since it is isomorphic to the additive group of $L(A)/N(A)$ we can equip it with a scalar multiplication borrowed from $L(A)/N(A)$. This latter space however is more conveniently represented as a direct summand of $L(A)$ as follows.

Let $M(A)$ denote the subset of $L(A)$ of functions satisfying

$$\sum\{f(a)|\mathrm{rank}(a) = i\} = 0$$

for each rank $i = 0, \ldots, n$. It is easy to verify that $M(A)$ is a subspace of $L(A)$ and that $L(A)$ is the direct sum

$$L(A) = M(A) + N(A)$$

so that $M(A)$ is isomorphic to $L(A)/N(A)$ in the usual way. For future use let $\mu_A\colon L(A) \to M(A)$ denote the projection on $M(A)$ along $N(A)$. The vector space operations on $Reg(A)$ are defined explicitly by:

$$p + q = \mathrm{reg}\,(\log p^{\mathrm{o}} + \log q^{\mathrm{o}})$$

$$k\,p = \mathrm{reg}\,(k \log p^{\mathrm{o}})$$

for all $p, q \in Reg(A)$ and all real numbers k. Then we have all that is needed to prove the following:

Proposition 17 *$Reg(A)$ is a real vector space with respect to the above operations and the restricted map $\mathrm{reg}\colon M(A) \to Reg(A)$ is an isomorphism of vector spaces.*

This shows us incidentally how to calculate the dimension of $Reg(A)$. Since L is the direct sum of M and N we have

$$\dim L = \dim M + \dim N.$$

But $\dim L = |A|$ whilst the dimension of N is just the number of ranks. Thus if A has m elements and ranks $0, \ldots, n$ then

$$\dim Reg(A) = m - (n+1).$$

We lose a dimension for each rank.

3.3 Belief Functions

In order to compare the theory of regular probability measures with the theory of belief functions we give an alternative description of regular measures in terms of their densities. The regular measures are those whose densities vanish on some element of every rank except the lowest, that is to say except on the top element.

Proposition 18 *Let A be a finite inflattice with $n+1$ ranks and let p be a proper probability measure on A with density m. Then p is regular if and only if for each $i = 1, \ldots, n$ there exists an element $a \in A$ of rank i such that $m(a) = 0$.*

The proof in one direction follows immediately from the construction of reg f in the proof of Lemma 11. The proof in the other direction is by induction on the sequence of functions $(m_i)_{i \leq n}$. Again we leave details to the reader.

For ease of comparison we restrict attention to belief functions that are proper in our sense, *i.e.* belief functions whose density on the top element is non-zero. Thus a *belief function* is a proper probability measure whose density satisfies $m(0) = 0$. We denote by $Blf(A)$ the set of belief functions on a finite inflattice A. Now the density of a regular probability measure must vanish at some point of every rank except the lowest, *i.e.* on the top element. Since the bottom element always occurs in a rank of its own it follows that every regular probability measure is a belief function.[6] Thus we have the inclusions

$$Reg(A) \subseteq Blf(A) \subseteq Pr(A).$$

The binary operations of combination for $Blf(A)$ and $Reg(A)$ are both based on the monoid operation in $Pr(A)$. But neither $Blf(A)$ nor $Reg(A)$ is closed under this operation. Some form of "renormalization" is needed in either case. Define the function $\beta\colon Pr(A) \to Blf(A)$ by

$$\beta(p)(a) = \frac{p(a) - p(0)}{p(1) - p(0)}$$

[6] We ignore the case of a free semilattice on the empty set.

for all $a \in A$. Then the binary operation on $\mathit{Blf}(A)$, which is called Dempster's rule of combination, is defined by

$$p + q = \beta(p \star q)$$

and $\beta \colon \mathit{Pr}(A) \to \mathit{Blf}(A)$ is a morphism of commutative monoids.

For the theory of regular probability measures on the other hand we have $\rho \colon \mathit{Pr}(A) \to \mathit{Reg}(A)$ defined by

$$\rho(p) = \mathrm{reg}\,(\log p^{\mathrm{o}}).$$

The binary operation on $\mathit{Reg}(A)$ is given by

$$p + q = \rho(p \star q)$$

and $\rho \colon \mathit{Pr}(A) \to \mathit{Reg}(A)$ is a morphism of commutative monoids. In fact both β and ρ are quotient maps and ρ factors uniquely through β.

3.4 Covariant Transformations

In an earlier section we defined the covariant probability functor Pr. If $f \colon A \to B$ is any morphism of finite inflattices and p is any proper probability measure on A then $\mathit{Pr}(f)(p) = (p^{\mathrm{o}} \circ f^{\mathrm{o}})^{\mathrm{o}}$ is a proper probability measure on B and $\mathit{Pr}(f)$ is a morphism of commutative monoids, but it does not follow that $\mathit{Pr}(f)(p)$ is regular whenever p is regular.

An obvious way of transforming $\mathit{Pr}(f)(p)$ into a regular measure is to apply the quotient map ρ on $\mathit{Pr}(B)$ which we refer to as ρ_B. This leads to the definition $\mathit{Reg}(f) = \rho_B \circ \mathit{Pr}(f)$. But this is not in general functorial. It does not follow that $\mathit{Reg}(g \circ f) = \mathit{Reg}(g) \circ \mathit{Reg}(f)$ whenever $g \colon B \to C$. We need to restrict to a special subclass of morphisms. What is needed is that $\mathit{Pr}(f)$ should map uniform measures on A to uniform measures on B. Then we can define $\mathit{Reg}(f)$ to be the unique morphism which makes the diagram

$$\begin{tikzcd}
\mathit{Pr}(A) \arrow[r, "\rho_A"] \arrow[d, "\mathit{Pr}(f)"'] & \mathit{Reg}(A) \arrow[d, dashed, "\mathit{Reg}(f)"] & A \arrow[d, "f"] \\
\mathit{Pr}(B) \arrow[r, "\rho_B"'] & \mathit{Reg}(B) & B
\end{tikzcd}$$

commute in the category of monoids. Such a morphism exists because the kernel of $\rho_B \circ \mathit{Pr}(f)$ then includes the kernel of ρ_A and uniqueness is obvious. Explicitly $\mathit{Reg}(f)$ is given by

$$\mathit{Reg}(f)(p) = \mathrm{reg}_B\, \log(p^{\mathrm{o}} \circ f^{\mathrm{o}}).$$

We postpone for the moment the proof that Reg is then functorial.

As an aside let us note the corresponding requirement for the theory of belief functions. Guided by the same idea that the above diagram should commute when we replace *Reg* by *Blf* and ρ by β, it is easily seen that an inflattice morphism f must then satisfy the additional requirement that $f(0) = 0$. Every inflattice epimorphism satisfies this condition but in general there are inflattice morphisms with $f(0) = 0$ for which $Pr(f)$ does not preserve uniformity, nor does the converse implication hold.

Now the general condition for $Pr(f)$ to preserve uniformity of measures is that f^{o} preserves equality of rank. Accordingly we restrict attention to such morphisms in the sequel. Let us denote by **rILf** the category whose objects are finite inflattices and whose morphisms are inflattice morphisms whose left adjoints preserve equality of rank. Composition is just composition of inflattice morphisms. This is a subcategory of the category of finite inflattices since the left adjoint of $g \circ f$ is $f^{\mathrm{o}} \circ g^{\mathrm{o}}$ and preservation of equality of rank is preserved under composition of left adjoints. We shall not attempt any investigation of the properties of this category beyond those that we need for immediate purposes.

We now give some examples of epimorphisms in this category.

Example 1 Let A be a finite inflattice and let $(PA)^{\mathrm{o}}$ be the power set of A ordered by opposite inclusion. This is the free inflattice on the underlying set of A. Then the map $\bigwedge\colon (PA)^{\mathrm{o}} \to A$ which sends a subset S of A to its meet $\bigwedge S$ preserves infima. (Since A is finite and the ordering is by opposite inclusion we need only verify that $\bigwedge(S \cup T) = (\bigwedge S) \wedge (\bigwedge T)$ and that $\bigwedge \emptyset = 1$.) In fact $\bigwedge$ is the map which exhibits an inflattice A as a canonical quotient of a free inflattice. Its left adjoint is the up-segment map which sends each $a \in A$ to $\uparrow(a)$ and by definition this preserves equality of rank.

Example 2 Let S be any upper set of a finite inflattice A and let $B = S \cup \{0\}$ be the union of S with the bottom element of A. Define $f\colon A \to B$ by

$$f(a) = \begin{cases} a & \text{if } a \in S \\ 0 & \text{otherwise.} \end{cases}$$

This is right adjoint to the inclusion of B in A so that f preserves infima and its left adjoint, the inclusion, preserves equality of rank. This is because S is an upper set so that the number of elements in B above a given element $b \in B$ is the same as the number of element above b in A, except possibly for the bottom element and this always occurs in a rank of its own. Thus B corresponds to a quotient of A in the category **rILf**.

Example 3 Suppose X is a finite set and that A is the free inflattice on X. Thus $A = (PX)^{\mathrm{o}}$ is the power set of X ordered by opposite inclusion. Let Y be a subset of X and let $B = \{S \subseteq X | Y \subseteq S\}$ be the collection of all supersets of Y. Then the map $(-) \cup Y\colon A \to B$ which sends a subset S of X to its union with Y is an inflattice morphism whose left adjoint preserves equality of rank. This is because the left adjoint is again the inclusion of subsets and the condition for equality of rank for sets in the quotient B is evidently (but not trivially) the same as it is for sets in the numerator A, namely equality of size.

More generally let b be some fixed element of a finite inflattice A and let $B = \downarrow(b)$. Then the more general situation corresponding to Example 3 is that of the quotient map $(-) \wedge b \colon A \to B$ which sends an element $a \in A$ to $a \wedge b \in B$. It must be pointed out, however, that $\downarrow(b)$ need not in general be a quotient object of A in the category **rILf** since the inclusion of $\downarrow(b)$ in A need not preserve equality of rank. Besides power sets (or their opposites), however, there is another significant case where this always occurs, namely when A is tree-like. We say that a finite inflattice A is *tree-like* if $\uparrow(b)$ is linearly ordered for all $b \in A$ except possibly $b = 0$. Then the inclusion of $\downarrow(b)$ in A always preserves equality of rank.

Examples 2 and 3 are related to the idea of conditionalisation in Boolean probability theory. First consider Example 3 in the case where $\downarrow(b)$ is indeed a quotient object of A. Applying *Reg* to the morphism $(-) \wedge b \colon A \to B$ yields a linear transformation from $\mathit{Reg}(A)$ to $\mathit{Reg}(B)$ corresponding to positive conditionalisation for inflattices. It tells us how to modify a regular measure on A if we discover that the proposition corresponding to the element b is true. The reason for this interpretation is that b, and anything which it entails, is mapped to the top element of the quotient inflattice B.

Example 2 relates similarly to negative conditionalisation, that is to say conditionalising on the information that one or more of the propositions represented in the the inflattice is false. This distinction between positive and negative conditionalisation is not needed in orthodox probability theory because of the assumed underlying Boolean structure. In general let A be a finite inflattice and suppose that we learn that each of the propositions in a subset R of A is false. Then we must identify all propositions, including the bottom element, which entail any of the (false) propositions in R. This amounts to forming a quotient of A in which each of the elements of R is mapped to the bottom element of the quotient. This is just the case of Example 2 when S is the complement of the downward closure of R.

Turning to monomorphisms we can restrict attention without loss of generality to inclusions. Then if B is a finite inflattice and A is a subset of B closed under infima, so that it is a sub-inflattice of B, the left adjoint to the inclusion of A in B preserves equality of rank only if A satisfies the following condition:

$$a \in A \text{ and } \mathrm{rank}(a) = \mathrm{rank}(b) \text{ imply } b \in A.$$

In other words if A includes any element of B of a given rank it must include all elements of that rank. In the power set case $B = PX$ this is satisfied if A includes X and all subsets of X of less than a given size. Then the left adjoint to the inclusion sends every subset in A to itself and every other subset to X and this preserves equality of rank.

Our next task is to show that *Reg* is functorial. To see this we make use of the isomorphism between $M(A)$ and $Reg(A)$. Recall that $M(A)$ is the subspace of $L(A)$ consisting of all real-valued functions on A whose sums across ranks separately vanish or, more significantly, whose arithmetic means across each rank separately vanish. Now let $f\colon A \to B$ be a morphism of finite inflattices whose left adjoint preserves equality of rank. First define $L(f)\colon L(A) \to L(B)$ by

$$L(f)(w)(b) = w(f^{\mathrm{o}}(b))$$

for all $w \in L(A)$ and $b \in B$. Then L is a functor from the category **rILf** to the category **R-VECf** of finite dimensional real vector spaces. Now define $M(f)\colon M(A) \to M(B)$ to be the unique linear transformation which makes the diagram

$$\begin{tikzcd} L(A) \arrow[r, "\mu_A"] \arrow[d, "L(f)"'] & M(A) \arrow[d, dashed, "M(f)"] & A \arrow[d, "f"] \\ L(B) \arrow[r, "\mu_B"'] & M(B) & B \end{tikzcd}$$

commute. Again it exists because the null space of $\mu_B \circ L(f)$ includes the null space of μ_A and uniqueness is obvious. Now we must show that M is indeed a functor from **rILf** to **R-VECf**. All we need show in effect is that $M(g \circ f) = M(g) \circ M(f)$ whenever $f\colon A \to B$ and $g\colon B \to C$ are morphisms of **rILf** and the details we leave to the reader. Essentially the reason is because if w belongs to $M(A)$ then the differences between the values of $M(g \circ f)(w)$ on any two elements of C of a given rank are the differences between the values of $w(f^{\mathrm{o}}(g^{\mathrm{o}}(c)))$ for elements $c \in C$ belonging to that rank. The condition that the arithmetic means across ranks should vanish then determines the actual function $M(g \circ f)(w)$ uniquely on each rank in terms of these differences. It is the same whether we pass via an intermediate relocation in $L(B)$ or whether we leave the entire relocation until we reach $L(C)$.

We now claim that for any morphism $f\colon A \to B$ of **rILf** the diagram

$$\begin{tikzcd} M(A) \arrow[r, "\mathrm{reg}_A"] \arrow[d, "M(f)"'] & Reg(A) \arrow[d, "Reg(f)"] & A \arrow[d, "f"] \\ M(B) \arrow[r, "\mathrm{reg}_B"'] & Reg(B) & B \end{tikzcd}$$

commutes. Again this is an exercise we leave to the reader. This fact together with a previous result establishes the following:

Proposition 19 *Reg is a functor from* **rILf** *to* **R-VECf** *and the linear transformations* reg_A *form the components of a natural isomorphism from* M *to Reg.*

3.5 Contravariant Transformations

The functor *Reg* is covariant, that is to say it transforms regular probability measures in the direction in which infima are preserved. Thus in the case of quotient maps it projects a regular measure on the numerator to a regular measure on the quotient. It is reasonable to ask whether there may also a well-behaved way of transforming regular measures in the opposite direction. In the case of quotient maps this would provide a way of lifting a regular measure to the numerator.

There is a practical reason for seeking such transformations which concerns the representation of conditional judgements. Suppose that a patient, who has been prescribed two drugs, is suffering some adverse condition and the question at issue is whether this is a case of an adverse drug reaction or whether there is some other explanation. An expert offers the following opinion on the basis of some particular item of evidence. "This tells me nothing directly about whether or not the condition is due to a drug. But it does suggest that if a drug is involved, it is more likely to be drug A than drug B."

Assuming that we can elicit a numerical expression of this opinion, we are being invited to represent it initially by a measure on the space whose top element corresponds to 'drug involvement'. The question is whether we can lift this measure on the quotient space to the whole space. There seems to be a need to do this. Another item of evidence may point directly and unconditionally to drug B, for example, while another may suggest, again unconditionally, that the patient was already predisposed to the condition for independent reasons. If we are to be able to combine together all these items of evidence it must be in a single space. Unless we can lift our first measure to the numerator, the evidence on which it was based will have to be disregarded even though it may discriminate strongly between the two drugs.

Exactly similar remarks hold in the case of judgements expressed conditionally on a certain proposition, or set of propositions, of the domain being false. The only difference lies in the type of quotient we have to consider.

There is an obvious way of mapping probability measures on an inflattice in a contravariant direction. If $f\colon A \to B$ is an inflattice morphism and p is a probability measure on B then Lemma 2 tells us that functional composition of p with f yields a probability measure on A. This was how we defined the covariant probability functor by composition of duals $p^\circ \circ f^\circ$. But whereas

$Pr(f)$ preserves the monoid operation, because we have taken duals before and after composing, direct composition with f does not. To see this we need only consider the quotient map $(-) \wedge b$ when $b \neq 1$. Composing with this will lift the unit of $Pr(\downarrow(b))$ not to the unit of $Pr(A)$ but to a measure concentrated on b (and therefore in fact an improper one). Clearly this is unsatisfactory for the purposes mentioned, and nothing can be done to rectify it. Suppose our expert were to say: “This tells me nothing directly about whether or not the condition is due to a drug. But it does suggest that if a drug is involved, it may as well be drug A as drug B.” The only plausible representation of this judgement is by the vacuous measure, whether on the quotient or the numerator.

The way in which we propose to deal with this question is based on the following fact:

Proposition 20 *If f is an epimorphism in* **rILf** *then $Reg(f)$ is a retraction in* **R-VECf**.

Proof Suppose $f\colon A \to B$ is an epimorphism. Define $Reg(f)^*$ for all $q \in Reg(B)$ by

$$Reg(f)^*(q) = \mathrm{reg}_A\,(\mu_B(\log q^{\mathrm{o}}) \circ f).$$

It must be verified that this is a linear transformation and that $Reg(f) \circ Reg(f)^*$ is the identity on $Reg(B)$. For both purposes we use the natural isomorphism between M and Reg. Then it is sufficient to prove that the map $M(f)^*$ defined for all $w \in M(B)$ by

$$M(f)^*(w) = \mu_A(w \circ f),$$

which is evidently a linear transformation from $M(B)$ to $M(A)$, is a right inverse for $M(f)$. To see this we observe that the differences between the values of $M(f)(M(f)^*(w))$ on any two elements of B of a given rank will be just the differences between the values of $w((f \circ f^{\mathrm{o}})(b))$ for elements $b \in B$ belonging to that rank with the final determination made uniquely by the condition of having vanishing arithmetic means across ranks. But if f is an epimorphism then $f \circ f^{\mathrm{o}}$ is the identity on B and the original function w already satisfied that condition. Hence $M(f)(M(f)^*(w)) = w$. The result now follows from the natural isomorphism between M and Reg. $\Box$

The proposal we adopt in the section on elicitation of opinions is that when an opinion is expressed conditionally on a quotient, it may be lifted to the numerator by $Reg(f)^*$ where f is the quotient map. This is a linear transformation so that, in particular, the vacuous measure on the quotient is lifted to the vacuous measure on the numerator and, furthermore, if any such lifted measure were then “conditionalised” back to the quotient we should have neither gained nor lost information.

It is interesting to note that, for certain classes of morphisms, Reg^* defined by $Reg^*(f) = Reg(f)^*$ is in fact functorial. This would be an important phenomenon to investigate in any development of these ideas. For the present we shall merely make practical use of the fact established without attempting a deeper understanding.

4 Some Philosophy

We have now developed to a certain extent the theory of regular probability measures on inflattices. Before turning to the parallel theory for suplattices, it will be useful to survey some of the ways suplattices and inflattices can arise in practice.

It is part of the underlying approach of this paper to consider each suitable algebraic structure as having its own peculiar concept of probability. The structures we have in mind are all partially ordered sets. So, for example, there is a separate notion of probability for Boolean algebras, distributive lattices, orthomodular lattices, semilattices etc. A given algebraic structure, however, often belongs, as an object, to a number of different categories. It is more correct therefore to say that each suitable category of algebraic structures has its own peculiar concept of probability. Every Boolean algebra, for example, is a semilattice and every free semilattice is a Boolean algebra. The concept of probability appropriate to such an object, when we use it to model a practical situation, depends on a prior assessment of the essential logical structure of the domain it is intended to model. We should not ask, for example, whether probabilities are additive in general, but whether it makes best sense to use a Boolean algebra (or distributive lattice) to model the logical relations between the propositions of that particular domain. If so, we must use additive measures. If not, we are entitled to employ another algebraic model with its own distinctive notion of probability.

Complete semilattices are particularly puzzling in this respect since every complete meet semilattice is also a complete join semilattice. Yet each has its own notion of probability just as each has its own concept of morphism. We must decide at the beginning, by considering the underlying logic of the domain, how we wish to view the object and that will simultaneously decide which is the appropriate concept of probability.

Consider the case of finite free semilattices. The free suplattice on a finite set X is the power set PX ordered by inclusion with the generators inserted by the singleton[7] map $\{-\}: X \to PX$. Suppose, for example, we are considering whom Paris will judge to be the fairest. The free suplattice generated by the set {Juno, Minerva, Venus} is shown in Figure 1. The element {Juno, Minerva} has the meaning of 'Juno or Minerva' and similarly for the other sets. We may believe the top element to be reasonably certain, *i.e.* that Paris will present the apple to one or other of them, but this representation does not oblige us to regard it as necessary.

[7]This means that if $f: X \to A$ is any function from X into the underlying set of a suplattice A there is a unique suplattice morphism $\overline{f}: PX \to A$ such that $\overline{f}(\{x\}) = f(x)$.

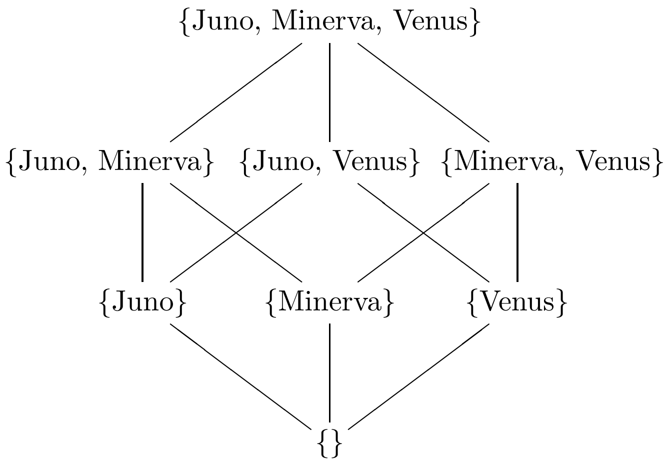

Figure 1: *The free suplattice generated by the set* {Juno, Minerva, Venus}.

The free inflattice on a finite set X, on the other hand, is again the power set of X with generators inserted as singletons, but this time the power set is ordered by opposite inclusion. Since a power set is isomorphic to its opposite via complements, we could take the free inflattice to be PX ordered by ordinary inclusion with the generators inserted as the complements of singletons, but little is gained by the inversion. In either case the insertion of generators map is universal up to isomorphism amongst functions from X into the underlying set of an inflattice.

Suppose I am considering why my car fails to start. I list the possible reasons in general as {fuel, ignition, mechanics}. The free inflattice generated by this set is shown in Figure 2. The element {fuel, ignition} now has

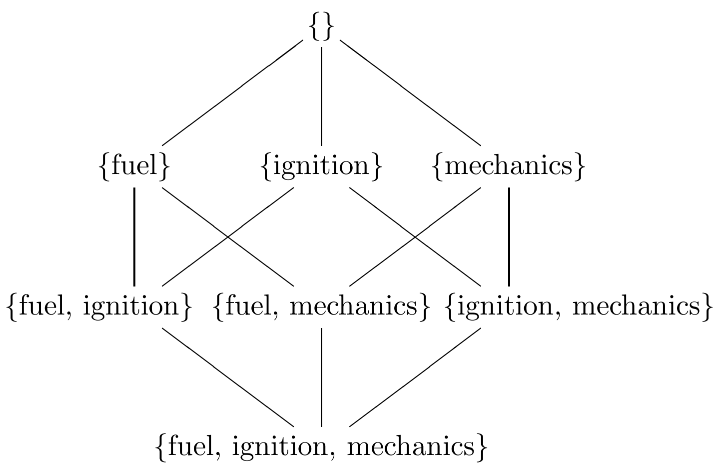

Figure 2: *The free inflattice generated by the set* {fuel, ignition, mechanics}.

the meaning of 'fuel and ignition' and similarly for the other sets. The bottom element ‘fuel and ignition and mechanics’ is unlikely, perhaps, but not impossible.

We see from this why it is intuitively as well as algebraically natural to use the ordering by opposite inclusion for the free inflattice. Whereas the free suplattice is generated as the joins of atoms, the free inflattice is generated as the meets of co-atoms. The sets involved, in either case, are just the sets of atoms or co-atoms making up the join or meet. The dual meanings of ‘or’ and ‘and’ provide the appropriate differences in meaning.

Whilst this helps to some extent to understand the differences between suplattices and inflattices, there are cases that it leaves unresolved. Suppose we are dealing with the free semilattice generated by just two elements. In that case the two elements that are neither top nor bottom are simultaneously atoms and co-atoms. Here is an example. Suppose that a patient is suffering from an adverse clinical condition and we are concerned whether this was caused by a specific drug the patient has been taking (the subject drug) or whether it has some other cause. We can represent the relation of entailment between these hypotheses by the partially ordered set of Figure 3 where a

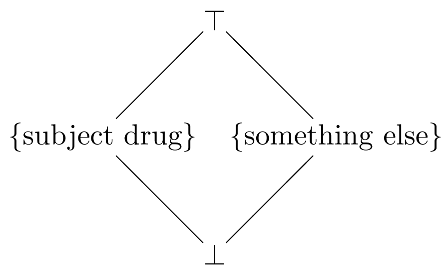

Figure 3: *A simple alternative.*

neutral notation has been used for the top and bottom elements so as not to prejudge the following question. Should we consider this to be an inflattice or a suplattice, or even a Boolean algebra? This will determine the appropriate notion of probability.

Setting aside the Boolean possibility for the moment, assume that this partially ordered set is to have a semilattice structure. Then the question hinges on which of two senses of cause we have in mind: whether we mean by cause a *necessary* condition or a *sufficient* condition. Here are two rough explications of the idea that the drug caused the event:

1. The event would not have occurred if the patient had not taken the subject drug.

2. The event only occurred because the patient took the subject drug.

In the first case taking the drug is viewed as a necessary condition for the event. But it is not said to be the only necessary condition. It might equally be true that the event would not have occurred if the patient's past medical history had not predisposed to the event, or if the patient had not been taking another drug at the same time (a drug interaction), etc. What we mean by (1) in simple terms is that "the drug had something to do with it", or that the drug was a partial cause. In the second case taking the drug is held to be a sufficient condition for the event. This would mean, literally, that all that was needed to cause the event was the taking of the drug. This seems puzzling unless it means elliptically that no other "abnormal" condition was necessary. But this unqualified way of speaking is used when we say "the whole cause of the trouble was the drug." And most would agree, to take an extreme example, that cyanide is a cause of death in just this sense.

The same dual interpretations must then be given to the alternative hypothesis that something else caused the event. In the first case it means that some other condition was a partial cause of the event. In the second case it means that some other condition was the whole cause of the event.

It is now clear which of these interpretations corresponds to the inflattice and which corresponds to the suplattice interpretation of the partially ordered set. If we are considering causation in the sense of a necessary condition we are dealing with an inflattice and if we are considering causation in the sense of a sufficient condition we are dealing with a suplattice. On the first interpretation the bottom element means that taking the drug was necessary and, at the same time, that something else was necessary, which is a meaningful contingent assertion—both may be true. In the second case it is the top element that corresponds to a meaningful contingent assertion, namely that either the drug was the whole cause or that something else was—both may be false.

It is up to the user to decide which of these two interpretations is suitable for a given enquiry. But if our concern is to monitor possible adverse effects of drugs, it seems that the necessary condition conception of causality is the appropriate one. To justify raising doubts about the safety of a drug, we need to establish with a sufficient degree of certainty that the adverse event would not have occurred if the drug had not been taken.[8] The situation appears to be different if we are considering the beneficial effects of drugs. For a wide range of drugs, few manufacturers would claim that the patient would not have improved if the drug had not been taken. The claim, rather, is that if the drug is taken an improvement will follow. The drug is recommended as a sufficient not a necessary condition for improvement—though even this may be more than a cautious manufacturer would be willing to claim.

Let us now consider whether, instead of either of these approaches, we should consider the previous partially ordered set to be a Boolean algebra. In that case the appropriate notion of probability will be the familiar additive concept, which in the present context we shall call the Boolean concept of probability, though in a common way of speaking it could equally be called the Bayesian concept of probability. A full discussion of this question lies beyond our present scope. We include a few remarks to indicate where the

[8] This is the position adopted in [3].

issues lie.

The classic domain in which additive probabilities over finite Boolean algebras is applied is that of games of chance terminating after a fixed finite time or number of moves. If we throw a die there are six specific outcomes. We can consider this as a space with six points. The classic duality between the category of finite sets and the category of finite Boolean algebras leads to the familiar correspondence between densities on the space and measures on the algebra. The general measure-theoretic approach to probability theory is ultimately only an attempt to extend this correspondence as far as possible consistently with the practical needs of the theory.[9] The essence of the approach, in either the discrete or continuous case, is to assume that the analysis begins with a space of points or most specific descriptions of possible outcomes.

We shall refer to this picture, based on a space of points, as the classical model. It has its roots in such simple phenomena as the throw of a die. It has been immensely successful for a wide range of phenomena. Classical statistical mechanics is perhaps its best known application. Here we assume that the system of interest is at any time in one and only one of a definite set of possible states. In principle we might be interested in the probability that the state of the system at any given time lies in an arbitrary subset of the state space. For technical reasons, however, it is fortunate that we can restrict attention to a subcollection of the set of all subsets of the state space. Ideally it should be enough to restrict ourselves to subsets definable in macroscopic terms such as pressure or volume, for example. The crucial point is that we assume that these subsets, or the propositions to which they correspond, have the structure of a Boolean algebra or sigma-algebra. And this assumption seems to be based in turn on the original idea that ultimately we are dealing with a space of points corresponding to the most specific possible descriptions of the state of the system. Were it not for technical difficulties, we should be willing in principle to attach a probability to any subset of this space and the corresponding algebra would be just the power set lattice considered as a Boolean algebra.

It is now clear that, despite its familiarity, this model is of limited applicability. Quantum mechanics has shown us that the classical picture is neither a necessity of thought, nor does it lead to correct experimental predictions. The propositions of orthodox quantum mechanics do not have the structure of a distributive lattice and the appropriate concept of probability is correspondingly different. It is now reasonable to conclude that no theory based on the model of a classical state space and a corresponding Boolean

[9] If the space is countable we can endow it with the discrete topology and the generalization is immediate if we assume countable additivity. If the space is not countable, a coarser topology is needed. It is usual then to take the associated algebra to be the algebra of clopen sets of the topology but there is no mathematical or practical necessity for doing so.

algebra could yield predictions consistent with experimental findings.[10]

What has this to do with our problem? It is not being suggested that the question of adverse drug reactions, for example, should be modeled by the projection lattice of a Hilbert space. The point is only that the Boolean model is not a necessity of thought. If we begin with a partial order representing an entailment relation between the propositions in some domain, any additional structure we assume needs to be justified positively on the basis of the specific empirical situation. In many cases it may be reasonable to impose this additional structure. It may be reasonable, for example, to picture a biological organism as similar to a classical physical system to the extent that at any time it is in one of a definite set of possible states. We may classify some of these states as "diseased" or as being diseased in some specific way. We may then be interested in the probability that the system is in one of these diseased states. In that case it may well be reasonable to suppose that we are dealing with a Boolean algebra of propositions whose corresponding concept of probability is the familiar Bayesian concept. On the other hand it may be an artificial assumption based only on a familiar pattern of thought and one we should do better without.

In summary, the use of Boolean probability theory carries with it an assumption about the "logic" appropriate to the underlying model. When we are dealing with causality, as in the example of adverse drug reactions, the Boolean assumption is not obviously the most appropriate. We have seen that the semilattice models permit a distinction between two senses of causality that appears to correspond to a distinction we are ready to make in practice. Until we have a more specific model of causality for such a domain, which may then not be a Boolean model, we could well do better to make only weak assumptions about the underlying logic of the domain.

We return now to the principal purpose of this paper which is to investigate the concept of probability under these weaker assumptions. We have dealt with the case of meet semilattices or inflattices. We turn now to the case of probability measures on join semilattices or suplattices.

5 Probability Measures on Suplattices

Definition 5 *A* probability measure *on a finite suplattice A is a real unit-interval valued function $p\colon A \to [0,1]$ satisfying*

$$p(\bigwedge S) + \sum_{R \subseteq S} (-1)^{|R|} p(\bigvee R) \leq 0$$

for every (finite) subset $S \subseteq A$.

Readers familiar with the theory of belief functions will recognise probability measures on suplattices as having the same mathematical structure as "upper probability" or "plausibility" functions.

[10] See [1] and [2].

Proposition 21 *Every probability measure on a finite suplattice A has a unique extension to a probability measure on the distributive lattice freely generated by A.*

The distributive lattice freely generated by a finite suplattice A is the lattice $(UA)^{\mathrm{o}}$ of all upper sets of A ordered by opposite inclusion. The insertion of generators is the up-segment mapping $\uparrow: A \to (UA)^{\mathrm{o}}$ which sends an element $a \in A$ to $\uparrow(a)$. To say a probability measure $\overline{p}$ on $(UA)^{\mathrm{o}}$ is an extension of a probability measure p on A is to say that $\overline{p}$ restricts to p along the map $\uparrow$. Thus $p(a) = \overline{p}(\uparrow(a))$.

From here we could develop the theory of probability measures on suplattices in parallel with the previous treatment of probability measures on inflattices— but independently. Such an approach would have the advantage of emphasizing that neither type of probability measure has any priority over the other. For the sake of economy, however, we shall make use of previous results together with the following relationship between probability measures on the two types of semilattice.

If A is a suplattice we denote by $A^{\sim}$ the opposite partially ordered set considered as an inflattice. (A^{o} would mean the opposite partially ordered set considered as a suplattice.) The meets of $A^{\sim}$ are just the joins of A. Now if p is any unit-interval valued function on A define $p^{\sim}$ by

$$p^{\sim}(a) = 1 - p(a)$$

for all $a \in A$.

Lemma 22 *p is a probability measure on the suplattice A if and only if $p^{\sim}$ is a probability measure on the inflattice $A^{\sim}$. Moreover the correspondence $p \mapsto p^{\sim}$ is bijective.*

Proof Apart from the definitions we need only the identity

$$\sum_{R \subseteq S} (-1)^{|R|} = \sum_{r \leq n} \binom{n}{r} (-1)^r = (1-1)^n = 0$$

where $n = |S|$. $\square$

We now digress briefly to discuss an interpretation of the relationship between a semilattice and its opposite which this result suggests. The elements of these semilattices are usually considered to be the same. All that changes is the ordering. Equally we can consider the change to occur in the mode of assertion. If an element a of a semilattice A is interpreted to mean 'it is true that P' for some proposition P, then the same element a of $A^{\sim}$ means 'it is false that P'. This does not mean, however, that there is any algebraic priority to A or its opposite. It is an external question which of the two modalities we consider applying to the elements of A and which to the elements of $A^{\sim}$.

This idea is made somewhat difficult to express by the fact that a proposition is normally considered to be an assertion rather than a denial. Another way of looking at the matter would be to consider the bare elements of the underlying set of a semilattice A as "states of affairs". From a certain point of view the order structure is the same for A as for $A^\sim$. All that has to be decided is the sense of the order and this corresponds to a decision whether the elements of the semilattice, now considered as propositions through their incorporation in the partial order, are to be taken in the mode of assertion or denial. Even this idea is difficult to express, however, on account of it being hard to conceive of states of affairs without the mediation of a propositional language. This then leads us back to the difficulty that propositions are normally considered to be assertions. Perhaps a better analogy would be the language of goals. It is easier to think of a goal as one that one might either aim to achieve or to avoid. With respect to the goal of checkmating my opponent in chess, I aim to achieve it and you aim to avoid it. Dually, my being checkmated is a goal I aim to avoid and you aim to achieve. Our goals are the same but our aims are opposite.

The purpose of this digression is to emphasise that the present theory incorporates the idea that since one should be able to consider any proposition as being either true or false, it should be possible to consider both the probability that the proposition is true as well as the probability that it is false. It is not necessary for that purpose, however, to assume the existence of complements, *i.e.* that for any proposition in a given domain there is another proposition in the same domain which is true if and only if the first proposition is false, and *vice versa.* In the present theory, if $p(a)$ is the probability that a is true then $1 - p(a)$ is the probability that a is false. Dually if $p(a)$ is the probability that a is false then $1-p(a)$ is the probability that a is true. But the propositions to which a probability is being attached cannot be regarded as belonging to the same domain. If one is considered to be an element of a suplattice, the other must be considered to be an element of the opposite inflattice and *vice versa.* Furthermore the formal structure of the corresponding probability measures is different in the two cases. But there is no internal way of telling which is the suplattice and which is the inflattice. That is an external decision which the user must make along the lines sketched in the preceding section.

We can now use Lemma 22 to transfer previous results to the case of probability measures on finite suplattices. The basic monoid operation is defined as follows. If p and q are probability measures on the finite suplattice A then

$$p \star q = (p^\sim \star q^\sim)^\sim.$$

The binary operation on the right hand side is the previously defined monoid operation for inflattice probability measures. It is clear then that the probability measures on a suplattice A form a commutative monoid under this new operation whose unit is the measure

$$v(a) = \begin{cases} 0 & \text{if } a = 0 \\ 1 & \text{otherwise.} \end{cases}$$

We can consider this as a default measure corresponding to a state of zero information or uninformative evidence. Thus contingent propositions in a suplattice have a default probability value of 1 whereas contingent propositions in an inflattice have a default probability value of 0. Propositions in a suplattice are assumed to be innocent, as it were, until proved guilty.

The covariant probability functor Pr is now defined as follows. (We use the same notation as for inflattices: the context will make it clear which functor is intended.) If A is a finite suplattice, let $Pr(A)$ denote the commutative monoid of probability measures on A. If $f\colon A \to B$ is a morphism of finite suplattices, let $f^\sim\colon A^\sim \to B^\sim$ denote the corresponding inflattice morphism between the opposite inflattices. Note that $f^\sim$ is the same as f as a function. Now define

$$Pr(f)(p) = Pr(f^\sim)(p^\sim)^\sim$$

where the functor Pr on the right is the inflattice functor. Then evidently we have:

Proposition 23 *Pr is a (covariant) functor from the category of finite suplattices to the category of commutative monoids.*

Just as in the case of inflattices we have a representation for probability measures on a finite suplattice in terms of their densities. This is just a restatement of Proposition 5.

Proposition 24 *A real-valued function p on a finite suplattice A is a probability measure on A if and only if there exists a real unit-interval valued function $m\colon A \to [0,1]$ with $\sum_{a\in A} m(a) = 1$ such that*

$$p(a) = 1 - \sum\{m(b)|a \leq b\} = \sum\{m(b)|a \not\leq b\}$$

for all $a \in A$. Moreover such a function, called the density *of p, is unique when it exists.*

The monoid operation has again a simple description in terms of this representation. If p has density m and q has density m then the density m of $p \star q$ is given by

$$m(a) = \sum\{m_p(b)m_q(c)|a = b \vee c\}.$$

The probability functor has the same simple description in terms of densities as before. If $f\colon A \to B$ is a morphism of finite suplattices and p is a probability measure with density m, then the density m_f of $Pr(f)(p)$ is given by

$$m_f(b) = \sum\{m(a)|f(a) = b\}.$$

6 Regular Measures on Suplattices

Our task here is to rework for suplattices our previous treatment of uniform measures on inflattices. We shall do this as rapidly as possible.

6.1 Uniform Measures

First we restrict attention to the submonoid of $Pr(A)$ of *proper* measures on a finite suplattice A. These are the measures for which

$$p(a) = 0 \text{ implies } a = 0$$

or in other words for which only the bottom element is assigned the minimum value of 0.

Evidently p is a proper probability measure on the suplattice A if and only if $p^{\sim}$ is a proper probability measure on the inflattice $A^{\sim}$. Again we shall change our terminology so that $Pr(A)$ now denotes just this submonoid which similarly satisfies the cancellation law. If p is a proper probability measure on A and $f\colon A \to B$ is a morphism of suplattices then $Pr(f)(p)$ is a proper probability measure on B. Thus Pr now denotes a functor from the category of finite suplattices to the category of commutative monoids satisfying the cancellation law.

We now define a proper probability measure p on a finite suplattice A to be *uniform* if and only if $p^{\sim}$ is uniform on $A^{\sim}$ and denote by $Un(A)$ the uniform probability measures on an suplattice A. The unit of the monoid $Pr(A)$ is always uniform as, for example, is the measure whose density is constant. Again this notion could be given an autonomous definition. The following proposition is now just a restatement of previous results.

Proposition 25 *If A is a finite suplattice then $Un(A)$ is an ideal in the commutative monoid $Pr(A)$ of all proper probability measures on A and $Pr(A)/Un(A)$ is an Abelian group.*

Again each equivalence class of $Pr(A)/Un(A)$ contains a distinguished representative. We introduce this concept by way of the corresponding inflattice concept.

Definition 6 *A proper measure on a finite suplattice A is* regular *if and only if $p^{\sim}$ is regular on the inflattice $A^{\sim}$.*

Evidently each element of $Pr(A)/Un(A)$ contains exactly one regular measure. Let $Reg(A)$ denote the regular measures on a finite suplattice A. We treat the vector space structure of $Reg(A)$ as follows.

For a finite suplattice A we say that two elements $a, b \in A$ *have the same rank* if and only if they have the same number of elements below them: thus if $|{\downarrow}(a)| = |{\downarrow}(b)|$. We now count the ranks from bottom to top by the integers 0,1, …and call this numerical assignment the *rank function.* For a finite power set suplattice PX, for example, with S a subset of X and $|X| = n$, rank(S) is just the cardinality of S with rank$(X) = n$ and rank$(\emptyset) = 0$. In general one can think of the rank function on a finite suplattice as a measure of the height of an element above the bottom, *i.e.* again a measure of the distance from the unit.

Now let $M(A)$ be the set of all real-valued functions f on a finite suplattice A satisfying

$$\sum\{f(a)|\mathrm{rank}(a) = i\} = 0$$

for each rank $i = 0, \ldots, n$. Evidently $M(A)$ is a real vector space under the pointwise operations.

We endow *Reg*(A) with a vector space structure by using previous work for inflattices. Thus we define

$$p + q = (\mathrm{reg}\,(\log p^{\sim\mathrm{o}} + \log q^{\sim\mathrm{o}}))^{\sim}$$

$$k\,p = (\mathrm{reg}\,(k \log p^{\sim\mathrm{o}}))^{\sim}$$

for all $p, q \in Reg(A)$ and all real k. Then we have the following:

Proposition 26 *Reg*(A) *is a real vector space with respect to the above operations which is isomorphic to* $M(A)$.

The isomorphism is given by the function which sends an element f of $M(A)$ to the element $(\mathrm{reg}\, f)^{\sim}$ of *Reg*(A).

6.2 Covariant Transformations

We saw previously that regular measures on finite inflattices are well behaved under transformations between their domains only if we restrict attention to a certain subcategory of the category of finite inflattices, specifically to morphisms whose left adjoints preserve equality of rank. A similar phenomenon occurs for regular measures on finite suplattices. The development of this part of the theory is now obvious. In line with the approach above we shall merely rely on the correspondence between p and $p^{\sim}$ and our previous results for regular measures on finite inflattices.

Let **rSLf** denote the category whose objects are finite suplattices and whose morphisms are suplattice morphisms whose right adjoints preserve equality of rank. Recall that two elements a and b of a suplattice A have the same rank if $|{\downarrow}(a)| = |{\downarrow}(b)|$. Again composition is just composition of suplattice morphisms. Now let $f\colon A \to B$ be a morphism in the category **rSLf**. Then for all $p \in Reg(A)$ we define $Reg(f)\colon Reg(A) \to Reg(B)$ by

$$Reg(f)(p) = (Reg(f^{\sim})(p^{\sim}))^{\sim}$$

where the functor *Reg* on the right is the previously defined inflattice functor. Thus *Reg* is clearly a functor from **rSLf** to the category of finite dimensional real vector spaces.

We have the following examples of epimorphisms in **rSLf** paralleling those of Section 3.4.

Example 1 Let A be a finite suplattice. Then $\bigvee\colon PA \to A$, which exhibits A as a quotient of a free suplattice, is an epimorphism in **rSLf**. Its right adjoint is the down-segment mapping.

Example 2 Let S be any lower set of a finite suplattice A and let $B = S \cup \{1\}$ be the union of S with the top element of A. Then $f\colon A \to B$ defined by

$$f(a) = \begin{cases} a & \text{if } a \in S \\ 1 & \text{otherwise.} \end{cases}$$

is a quotient map in **rSLf**. Its right adjoint is the inclusion.

Example 3 Suppose that $A = PX$ is a finite power set suplattice, ordered by inclusion, and that Y is some subset of X. Let $B = \{S \subseteq X | Y \subseteq S\}$ be the collection of all supersets of Y. B is closed under intersections, namely the meets of A, and therefore corresponds to a quotient suplattice. The quotient map is $(-) \cup Y\colon A \to B$ which sends a subset $S \subseteq X$ to its union with Y. Its right adjoint is again the inclusion of subsets and this preserves equality of rank since two supersets of S include the same number of sub-supersets of S if and only if they are of the same size. If we apply the functor *Reg* to this quotient map we obtain a linear transformation from $\mathit{Reg}(A)$ to $\mathit{Reg}(B)$ which corresponds to negative conditionalisation for suplattices. It tells us how to modify a regular measures on $A = PX$ if we discover that the proposition corresponding to a certain subset S is false. This is because S, and anything which entails it, becomes the bottom element of the quotient suplattice.

For a general suplattice A the corresponding picture is that of projecting $\mathit{Reg}(A)$ onto $\mathit{Reg}(\uparrow(b))$, for some element $b \in A$, along the quotient map which sends an element $a \in A$ to $a \vee b$. But again it must be pointed out that in general $\uparrow(b)$ need not be a quotient object of A in **rSLf**.

For suplattices it is Example 2 that corresponds to positive conditionalisation, that is to say conditionalising on the information that one or more of the propositions represented in the suplattice is true. In general let A be a finite suplattice and suppose that we learn that each of the propositions in a subset $R \subseteq A$ is true. Then we must identify all propositions, including the top element, which are entailed by any of the (true) propositions in R. This amounts to forming a quotient of A in which each of the elements of R is mapped to the top element of the quotient. This is just the case of Example 2 when S is the complement of the upward closure of R.

Turning to monomorphisms, let A be a sub-suplattice of B. In other words let A be a subset of B which is closed under the suprema of B. Then the right adjoint to the inclusion of A in B preserves equality of suplattice rank only if A satisfies the closure condition:

$$a \in A \text{ and } \operatorname{rank}(a) = \operatorname{rank}(b) \text{ implies } b \in A.$$

In the power set case $B = PX$ this means that A must include the empty set and all subsets of X of greater than a given size. Then the right adjoint to the inclusion sends every subset in A to itself and every other subset to the empty set. Clearly this preserves equality of rank.

6.3 Contravariant Transformations

Our discussion here parallels exactly the discussion of contravariant transformations for probability measures on inflattices. Again we see the immediate practical use of such transformations in expressing the effect of conditional opinions. Consider again whether Paris is likely to choose Juno, Minerva or Venus. We can imagine an acquaintance expressing the following opinion based on knowledge of a certain trait of Paris's character. "This tells me nothing directly about which of the three Paris is likely to choose. But it does suggest that if he is not going to choose Venus, he is more likely to choose Minerva[11] than Juno." This is a negative conditional judgement, expressed conditionally on a certain proposition being false. Given a numerical expression of this opinion, we can represent it by a measure on the quotient space whose bottom element is 'Venus'. It must then be lifted to a measure on the whole space. The other situation is one of a positive conditional judgement. Our informant might have said: "This tells me nothing directly about which of the three Paris is likely to choose. But it does suggest that if he is going to choose either Juno or Minerva he is more likely to choose Minerva." Again we must lift this conditional judgement to the numerator. The only difference lies in the type of quotient on which the measure is first defined.

Our solution to this question is again based on the following:

Proposition 27 *If f is an epimorphism in* **rSLf** *then $Reg(f)$ is a retraction in* **R-VECf**.

Reg is of course the suplattice probability functor and the proposition is just a restatement of the corresponding inflattice result. The right inverse $Reg(f)_*$ of $Reg(f)$ that we shall use for the purpose of lifting a quotient measure to the numerator is defined as follows. Suppose $f\colon A \to B$ is a suplattice morphism whose right adjoint preserves (suplattice) rank and that $q \in Reg(B)$. Then explicitly

$$Reg(f)_*(q) = (\mathrm{reg}_{A^\sim}(\mu_{B^\sim}(\log q^{\sim\mathrm{o}}) \circ f^\sim))^\sim.$$

Again the significance of this being a right inverse for $Reg(f)$ is that any measure lifted by this map to a numerator and then conditionalised back to the quotient is the same measure on the quotient as we began with.

7 Independence

Before making practical proposals for constructing regular measures in the next section, we must say something about the concept of independence. Following the original insight of Shafer [11], we intend to interpret addition

[11] Recall that Minerva was offering glory and Juno riches.

of regular measures as the mathematical operation corresponding to combination of independent items of evidence. If we have two items of evidence and regular measures p and q expressing the quantitative impact of each separately, the impact of the combined evidence is to be expressed by the (vector) sum $p + q$ provided they are independent. But what does it mean to say that two items of evidence are independent?

Suppose two witnesses both identify Jones as the person seen running away from the scene of the crime. Does their combined testimony provide more support for Jones being the culprit than either would on its own? It depends on whether or not they are independent. If our two witnesses are Mrs Archer, who lives opposite, and a passing cyclist who is unacquainted with Mrs Archer, we might be willing to accept their testimonies as independent. Their combined testimonies will count more than either on its own. But if our second witness is Mrs Archer's nextdoor neighbour Mrs Baker, and we know they have discussed little else since the incident, we might be inclined to rate their combined testimony as worth little more than the average of the two. Or again, suppose we are concerned with the voltage across a car battery. If two separate voltmeters both show a satisfactory voltage, this is more reassuring than a satisfactory reading from a single voltmeter, or even than two measurements from the same instrument.

It is important to emphasise that this intuitive concept of independence is prior to any reconstruction of it in a formal theory. The idea of independent testimony, for instance, must be as old as jurisprudence itself. Different theories can only be more or less successful at explicating it.

In the additive theory of probability on Boolean algebras, an event A is said to be "independent" of an event B if the conditional probability $P(A|B)$ coincides with the unconditional probability $P(A)$. Learning that B had occurred would have no effect on the probability of A. Independence is therefore a relation between events, rather than bodies of evidence, and in fact a symmetric one since the condition for independence of A and B is equivalent to the symmetric product rule

$$P(A\&B) = P(A)P(B).$$

Applying the product rule interpretation to the independence of evidence requires some care. One possibility, in the case of measuring instruments for example, is to say that the readings of two measuring instruments form independent items of evidence if the reliability of either instrument is independent, in the sense of the product rule, of the reliability of the other. On this interpretation the product rule is being applied to propositions about the reliability of the instruments, at a higher level as it were, and not to propositions at the base level about the state of affairs with respect to the battery itself. Attempting to use the product rule at the base level to explicate this notion of independence usually leads to unhelpful or irrelevant complications.

It is common, in the additive theory, to speak as though the independence of events were a factual question. This is true if the probability measure is given. But is the question which probability measure is appropriate to a given situation itself a factual one? Possibly so, if we all agree that probabilities are to be determined by counting frequencies and if we agree on what the relevant frequencies are. But we are not always so fortunate. Generally speaking there is no algorithm for determining the correct probability measure and there is therefore no algorithm for deciding whether or not two events are independent. We normally proceed by first making an intuitive judgement whether certain events are independent and, if so, we use the theoretical explication of the idea of independence as a tool to extend the domain over which probabilities are defined. Until there are reasons for rejecting it we are content to use this extension as a provisional measure.

The same applies to the theory of probability measures on semilattices. We say that two items of evidence are independent—in the judgement of some individual—if their joint impact is satisfactorily expressed for that individual by $p + q$, where p and q relate to the two items separately. This appears to involve a circularity. At the beginning we said that the impact of two items of evidence is expressed by $p + q$ if they are independent. Now we are saying that they are independent if $p + q$ expresses their joint impact. The solution is the same as before. If two items of evidence are judged to be intuitively independent, there becomes a presumption in favour of $p + q$ as the correct expression for their joint impact. Only if there is a positive reason for rejecting it should we look for another measure, either by analysing the problem further into simpler components or by direct assessment of the joint impact of the two items. In this way we have, if nothing else, at least a default measure to fall back on until another method presents itself.

This procedure relies heavily on the intuitive notion of independence. We need to develop ways of thinking that will help us sharpen the concept. We are helped to some extent by the fact that evidence does not always, or even often, present itself unsolicited. We need to ask relevant questions. Our task will then be made easier if we can analyse the problem into independent factors. As an example consider the problem of assessing possible adverse drug reactions. An examination by Castle *et al* [3] identified the following factors as relevant to the question whether an adverse clinical manifestation in a particular patient should be interpreted as a case of an adverse drug reaction:

1. When the event occurs in clinical practice, how usual it is that it is drug induced.

2. Whether or not there was a positive rechallenge, *i.e.* whether the problem recurred when the drug was reintroduced.

3. Whether the site of drug administration was compatible with the event.

4. Whether the time to onset of the event in the particular patient was compatible with a drug related effect.

5. Whether any of the diseases which the patient has can in themselves lead to the event.

6. Whether there is a pharmacological reason why the suspect drug might cause the particular event.

7. The frequency of occurrence of the event in normal practice independent of the particular patient.

8. Whether or not there was a positive dechallenge, *i.e.* whether the problem resolved when the drug was stopped.

9. Whether the particular patient has had the particular event previously in his or her life.

10. Whether the particular event was a recognized adverse reaction of any of the other drugs taken by the patient, as acknowledged by the different manufacturers.

11. Whether in fact the time to onset of the event is appropriate to any of the other drugs being taken by the particular patient.

12. Whether the manufacturer acknowledges that the event may be an adverse reaction to the suspect drug.

13. Whether the dose of the suspect drug was appropriate for the particular patient.

14. Whether the patient has any other characteristics which could predispose to the event, *e.g.* smoking, obesity, contraceptive method, age, *etc.*

It is not easy to assess whether these factors are independent though there is a reasonable presumption that many pairs are. However there is one error we must guard against. Suppose, for example, there is a known pharmacological reason why the subject drug might cause the adverse event and also that the time to onset of the event after the subject drug was introduced was almost exactly what one would expect if it were due to the drug. On one interpretation these are certainly not independent. If there is a pharmacological reason why the drug might cause the event, there is presumably also a pharmacological reason why it would cause the event more or less when we would expect if it were due to the drug. But this is not the question at issue. The question is rather whether, if we knew both that the pharmacology was right and that the time to onset was right, would we then have more evidence for drug involvement than if we only knew either on its own? On the face of it, yes. The question is only, by how much? Is it like

Mrs Archer and the passing cyclist, or is it like Mrs Archer and her neighbour Mrs Baker, or is it somewhere in between?

There is no easy way to settle these questions. The points to make are

- that the intuitive concept of independence of evidence is a primitive part of our common inductive intuition
- that there is so far no algorithm to substitute for individual judgment in deciding when two bodies of evidence are independent
- that the present theory is in no worse a position in this respect than the Bayesian theory.

The reader will find somewhat similar views expressed in [14].

8 Elicitation

In a given practical situation, how can we construct a regular measure that expresses the impact of a given body of evidence? We consider first the case of inflattices.

In the theory of belief functions a special role is played by the "simple support functions". Besides the vacuous measure, these are belief functions on a finite power set PX whose densities are given, for some fixed non-empty proper subset A of X and some number s in the real unit interval, by

$$m(B) = \begin{cases} 1-s & \text{if } B = X \\ s & \text{if } B = A \\ 0 & \text{otherwise.} \end{cases}$$

These are understood to express evidence that points precisely and unambiguously to a single non-empty subset A of X ([11, p.75]).

In the case of power sets, in particular, these measures will play a similar basic role in the present theory. Each such simple support function is a regular measure and in fact every regular measure can be expressed as a linear combination of such measures. We adopt the proposal that when we are dealing with a free inflattice A, an item of evidence that weighs precisely in favour of a single element $a \in A$, to a given degree s in the open unit interval, is to be represented by the measure

$$p(b) = \begin{cases} 1 & \text{if } b = 1 \\ s & \text{if } a \leq b \neq 1 \\ 0 & \text{otherwise} \end{cases}$$

where is it is assumed that a is neither 0 nor 1. Now suppose that an item of evidence weighs precisely *against* a certain non-zero element $a \in A$ to a given degree. We have already discussed the way this is viewed in the theory of belief functions and reasons for taking another approach. Our view is based on the idea that two items of evidence, one of which weighs precisely in favour of a proposition to a given degree and the other of which weighs precisely against that proposition to the same degree, should cancel out when combined. This forces us to deal with contrary evidence as follows. If an item of evidence weighs precisely against an element a to degree s, its effect is expressed by the additive inverse in the group of regular measures of the simple measure that would express the opposite situation in which the evidence weighs precisely in favour of a to the same absolute numerical degree.

Consider again the example of why the car fails to start. Suppose we have an item of evidence that weighs to degree 0.4 precisely in favour of there being a fault in the ignition system. This leads to the measure of Table 1. If the evidence had weighed precisely against there being a fault in

(‘’, 1)
(‘fuel’, 0)
(‘ignition’, 0.4)
(‘mechanics’, 0)
(‘fuel and ignition’, 0)
(‘fuel and mechanics’, 0)
(‘ignition and mechanics’, 0)
(‘fuel and ignition and mechanics’, 0)

Table 1: *Evidence in favour of* ‘ignition’.

the ignition system to the same numerical degree 0.4 then, according to our proposal, this would lead to the additive inverse of the preceding measure which is given in Table 2.

(‘’, 1)
(‘fuel’, 0.4)
(‘ignition’, 0)
(‘mechanics’, 0.4)
(‘fuel and ignition’, 0)
(‘fuel and mechanics’, 0.16)
(‘ignition and mechanics’, 0)
(‘fuel and ignition and mechanics’, 0)

Table 2: *Evidence against* ‘ignition’.

If we thought that something else might also be wrong, evidence to degree 0.4 against ‘ignition’ would lead to the measure of Table 3 which shows the general pattern for the effect of evidence weighing against a generator of a free inflattice.

We observe that the free semilattice on three generators is in fact a Boolean algebra and the complement of ‘ignition’ in this lattice is ‘fuel and

(‘’, 1)
(‘fuel’, 0.4)
(‘ignition’, 0)
(‘mechanics’, 0.4)
(‘something else’, 0.4)
(‘fuel and ignition’, 0)
(‘fuel and mechanics’, 0.16)
(‘fuel and something else’, 0.16)
(‘ignition and mechanics’, 0)
(‘ignition and something else’, 0)
(‘mechanics and something else’, 0.16)
(‘fuel and ignition and mechanics’, 0)
(‘fuel and ignition and something else’, 0)
(‘fuel and mechanics and something else’, 0.064)
(‘ignition and mechanics and something else’, 0)
(‘fuel and ignition and mechanics and something else’, 0)

Table 3: *Evidence against* ‘ignition’.

mechanics’. If we assumed that evidence against ‘ignition’ is equivalent to evidence of the same strength in favour of its complement, this would lead to the measure of Table 4. This seems less satisfactory insofar as evidence

(‘’, 1)
(‘fuel’, 0.4)
(‘ignition’, 0)
(‘mechanics’, 0.4)
(‘fuel and ignition’, 0)
(‘fuel and mechanics’, 0.4)
(‘ignition and mechanics’, 0)
(‘fuel and ignition and mechanics’, 0)

Table 4: *Evidence in favour of* ‘fuel and mechanics’.

against there being a fault in the ignition system ought not to support the assertion that there is a fault in both of the other two systems as strongly as it supports either of the two weaker assertions.

Consider now the dual case of a suplattice. The simplest situation is one in which we have evidence that weighs precisely *against* a given proposition to a specified degree. Suppose we are dealing with a free suplattice A (a power set) and that the evidence weighs precisely against an element $a \in A$ to degree s. Then we propose to represent such an item of evidence by the measure

$$p(b) = \begin{cases} 0 & \text{if } b = 0 \\ 1 - s & \text{if } 0 \neq b \leq a \\ 1 & \text{otherwise} \end{cases}$$

where again it is assumed that a is neither 0 nor 1.

Now suppose that an item of evidence weighs precisely in favour of an element a to a given degree s. This will be represented by the additive inverse in the group of regular (suplattice) measures of the measure corresponding to evidence against a to the same degree.

Here are some examples. Suppose again we are considering the judgement of Paris and that there is evidence of strength 0.4 against his choosing Venus. This leads to Table 5. On the other hand evidence in favour of 'Venus' to the

('Juno or Minerva or Venus', 1)
('Juno or Minerva', 1)
('Juno or Venus', 1)
('Minerva or Venus', 1)
('Juno', 1)
('Minerva', 1)
('Venus', 0.6)
('', 0)

Table 5: *Evidence against* 'Venus'.

same numerical degree of 0.4 leads to its additive inverse which is given in Table 6. Any proposition including 'Venus' as a disjunct has probability 1.

('Juno or Minerva or Venus', 1)
('Juno or Minerva', 0.84)
('Juno or Venus', 1)
('Minerva or Venus', 1)
('Juno', 0.6)
('Minerva', 0.6)
('Venus', 1)
('', 0)

Table 6: *Evidence in favour of* 'Venus'.

This is as it should be. Every proposition, except for the bottom element, has a default probability value of 1. The impact of evidence in favour of 'Venus' can only be expressed by a reduction in the probability of alternatives.

Now consider conditional judgements. The general theory was discussed in Sections 3.5 and 6.3. We pointed out that the relevant quotient algebras differ according to whether the conditional judgement is made positively or negatively, and the distinction goes in different directions for suplattices and inflattices. For illustration we restrict ourselves to suplattices. Suppose, to provide more scope for the example, that Discord has offered Paris the additional choice of awarding the apple to Diana and that there is evidence of strength 0.4 in favour of his choosing either Juno or Minerva conditional on his not choosing Diana. Then according to our proposal this determines in the first place a measure on the quotient space with bottom element ‘Diana’, which is then lifted to the numerator. The resulting measure is that of Table 7. ‘Juno’ and ‘Minerva’ become individually more probable than ‘Diana’

(‘Diana or Juno or Minerva or Venus’, 1)
(‘Diana or Juno or Minerva’, 1)
(‘Diana or Juno or Venus’, 0.8884)
(‘Diana or Minerva or Venus’, 0.8884)
(‘Juno or Minerva or Venus’, 0.9560)
(‘Diana or Juno’, 0.8046)
(‘Diana or Minerva’, 0.8046)
(‘Diana or Venus’, 0.6930)
(‘Juno or Minerva’, 0.9560)
(‘Juno or Venus’, 0.7068)
(‘Minerva or Venus’, 0.7068)
(‘Diana’, 0.5255)
(‘Juno’, 0.6231)
(‘Minerva’, 0.6231)
(‘Venus’, 0.3738)
(‘’, 0)

Table 7: *Evidence in favour of* ‘Juno or Minerva’ *conditional on* ‘Diana’ *being false.*

whilst ‘Venus’ becomes less probable.

It may seem surprising that a judgement made only conditionally on Paris not choosing Diana, and therefore apparently neutral as to whether he will or will not, should lead to a measure assigning a specific intermediate value of 0.5255 to this outcome. But this probability ought not to be 0, since it is not asserted that Paris will not choose Diana, nor should it be 1 since such a measure would unable to discriminate between any two propositions which included ‘Diana’ as a disjunct, since both would have probability 1. Since there is no intermediate value which intuition alone suggests we may as well be guided, at least provisionally, by the theory.

Now consider the case of a positive conditional judgement. Suppose we had evidence of the same numerical strength 0.4 in favour of Paris choosing either Juno or Minerva, conditional on his choosing either Juno or Minerva or Venus. This would lead to the measure of Table 8. The positive assumption that Paris will choose Juno or Minerva or Venus has a greater impact than the negative assumption that he will not choose Diana.

In dealing with this case we first form a measure on the appropriate quotient and then lift it to the numerator. The quotient suplattice here corresponds to the subset of the power set lattice from which ‘Juno or Minerva or Venus’ has been removed. It has been mapped to the top element. Since this is no longer a free suplattice, our earlier remarks about how to deal with

('Diana or Juno or Minerva or Venus', 1)
('Diana or Juno or Minerva', 1)
('Diana or Juno or Venus', 0.8741)
('Diana or Minerva or Venus', 0.8741)
('Juno or Minerva or Venus', 0.8659)
('Diana or Juno', 0.7796)
('Diana or Minerva', 0.7796)
('Diana or Venus', 0.6537)
('Juno or Minerva', 0.8659)
('Juno or Venus', 0.6455)
('Minerva or Venus', 0.6455)
('Diana', 0.4647)
('Juno', 0.5510)
('Minerva', 0.5510)
('Venus', 0.3306)
('', 0)

Table 8: *Evidence in favour of* 'Juno or Minerva' *conditional on* 'Juno or Minerva or Venus' *being true.*

simple unconditional judgements, which were restricted to free semilattices, are incomplete. There is nonetheless an obvious way of extending the idea of a simple regular measure to deal with this situation and it has been employed here. It has also been implemented in the program to be found in the Appendix. Also discernible there are points of contacts with the concept of "the weight of evidence" of [11, Ch.5]. There is not sufficient space here to deal fully with these ideas. We leave them at present to be investigated by the interested reader.

9 Further Developments

An obvious omission from our discussion of general probability measures on inflattices and suplattices is a treatment of measures on tensor products. The free suplattices, for example, on sets X and Y are the power sets PX and PY and their suplattice tensor product is the power set of their cartesian product, namely $P(X \times Y)$, with the universal bimorphism sending (S,T) to $S \times T$. A dual result holds for inflattices. It is clearly important to investigate how probability measures on these various semilattices are related. (Compare [12].)

Closely related is the question of probability measures on Homset semilattices. If A and B are suplattices, the set $Hom(A,\ B)$ of all suplattice homomorphisms from A to B is itself a suplattice under the obvious coordinatewise partial ordering, and similarly for inflattices. Probability measures on these Homsets can be thought of as "random homomorphisms" of the appropriate type from A to B. This is essentially the same topic as the previous one in view of the natural isomorphisms[12]

$$Hom(A,\, B) \cong (A^{\mathrm{o}} \otimes B)^{\mathrm{o}}$$

$$A \otimes B \cong Hom(A,\, B^{\mathrm{o}})^{\mathrm{o}}$$

Turning to the theory of regular measures, we need to know more about the categories **rILf** and **rSLf**. For example, should we in general be dealing with certain subcategories of these categories to have a really smooth running theory in the finite case? And what happens when we relax the condition of finiteness? Is the idea of 'preservation of rank' necessarily a discrete combinatorial idea or is there a suitable generalisation? It seems significant in this connection (1) that what we have described as 'uniform' measures are just the symmetric measures in the case of free semilattices, *i.e.* those that are invariant under all automorphisms, and (2) that in the finite case the morphism which exhibits a semilattice as a canonical quotient of a free semilattice (of the appropriate type) has an adjoint (left or right) which preserves equality of rank.

Apart from these technical questions, the most urgent need is for a decision theory to direct the practical application of these ideas.

Appendix

The appendix describes a program implementing the proposals of the paper. It is written in the functional programming language Standard ML which is described in the second section.

The Program

The purpose of the program is to define operations of vector addition and scalar multiplication for regular measures on semilattices. It also defines functions `supcon` and `infcon` which take a semilattice and a set of data and return the regular measure on the semilattice constructed from the data in accordance with the proposals of Section 8.

The visible part of a measure is represented by a list of pairs of the form (a, x), where a is an element of the lattice and x is a number in the real unit-interval, the value of the measure on that element. There are many ways in which the elements of a lattice and the order relation between them can be represented. For simplicity we have chosen to represent the elements as elements of a power set or its opposite. The order relation between elements is then to a large extent already encoded in the representation of the elements themselves. In general it is assumed that the semilattice is presented either

[12] See [9, pages 5,6].

as a sub-suplattice of a power set ordered by inclusion or, equivalently, as a sub-inflattice of a power set ordered by opposite inclusion.

Sets are represented by lists of booleans. Thus the subset $\{x, z\}$ of the base set $\{x, y, z\}$ is represented by `[true,false,true]`. In the programming language used below this will be an object of type: `bool list`. The infix operators `C` and `U` defined in the program then correspond to the relation of inclusion and the operation of union respectively. But the reader should always bear in mind that the ordering is by opposite inclusion for inflattices so that `a C b` then means that the object denoted by `b` is below that denoted by `a`.

It is assumed that the semilattice comes already ranked in the sense appropriate to either suplattices or inflattices. For free semilattices this will just be ranking according to size. The free semilattice on three generators is then given by

```
[[[true,true,true]],
[[true,true,false],[true,false,true],[false,true,true]],
[[true,false,false],[false,true,false],[false,false,true]],
[[false,false,false]]]
```

of type: `bool list list list`. It is assumed that the rank containing just the top element comes first in the case of suplattices, and the rank containing just the bottom element comes first for inflattices. This means that `[[true,true,true]]` comes first in either case for the semilattice generated by a 3-element set. The ranks are then presented in order of decreasing rank according to the appropriate definition of rank. The order of presentation of elements within a given rank is unimportant.

Measures are constructed in response to data inputs of the type discussed in the text. A datum is a pair $((b, (pos, neg)), s)$ where b is an element of the semilattice, *pos* and *neg* are lists of elements of the semilattice and s is a real number in the open interval $(-1, +1)$. A datum corresponds to the judgement that an item of evidence bears precisely on b to degree s conditional on the propositions in the list *pos* being true and the propositions in the list *neg* being false. Positive values of s correspond to items of evidence weighing in favour of b and negative values correspond to items of evidence weighing against b. In the present representation a datum has type:

```
(bool list * (bool list list * bool list list)) * real.
```

A list of data is converted by an operation called `profile` into a function on the semilattice whose values across ranks sum to zero. This is an element of the space $M(A)$ in the notation of Sections 3.2 and 6.1. The operation `regularise` is then applied in the sense appropriate to the type of semilattice to yield the corresponding regular measure. The correspondence is given by the natural isomorphism of Proposition 19 or by the corresponding isomorphism for suplattices.

Since the vector space operations on regular measures correspond to the more directly computed vector space operations on their profiles, it would be wasteful to throw away the profile of a regular measure once it has been computed. We therefore define an abstract data type `MEASURE` whose objects include both elements. A measure now consists of this pair together with an indicator of type `SENSE` to show whether it is a supmeasure or an infmeasure. The profile of a measure is hidden from view within the abstract data type. The regular measure itself can be recovered by the operation `find`.

The operations `supcon` and `infcon` assume that the input data correspond to independent items of evidence. If they correspond to assessments of the same item of evidence by N different experts, the averaged measure for this item is obtained by applying `supcon` or `infcon` followed by a scalar multiplication by the factor $1/N$, assuming the opinions are to be given equal weight.

The operation of vector addition `++` is defined as a binary infix operation. But if we wanted to combine three or more regular measures it would be more efficient to add their profiles first before applying the operation `reg`. It is simple to define such an operation of addition for any finite number of arguments so that `reg` is only applied once. A similar observation applies to the operation of scalar multiplication `**` for forming averages, especially with unequal weights.

Although some attention has been paid to questions of efficiency, greater attention has been paid to clarity of mathematical content. The operation `profile` in particular could be made more efficient. In general it is clear that for a single vector addition, for example, it is the operation `reg` which consumes the bulk of the time. For a free semilattice on n generators this increases exponentially with n, assuming sequential processing, with the base of the exponential lying between 3 and 4 depending on the efficiency of the test for the order relation between elements relative to the efficiency of real addition. For larger values of n approximation techniques are needed. Occurrences in the present code of the conditions `x = 0.0` and `z = 0.0` show where such techniques can be applied.

Whatever the computational merits of representing a lattice as an object of type: `bool list list list`, some other representation is certainly more convenient for intelligible input and output of data. Another way of representing a free semilattice has been indicated in Section 8. More extended programs involving alternative representations and making use of the modules facility of ML are available from the author on request.

The Language

The language chosen for the following program is Standard ML.[13] Some comments may assist readers to whom it is unfamiliar. For more information see [6] and [16].

[13] ML stands for 'metalanguage'.

SML is a strongly typed functional programming language. Each legal expression has a type which is determined automatically by the compiler. The basic types employed in the present program are booleans and reals. More complex types are obtained by forming lists of objects of a given type (list types), n-tuples of objects of given types (product types) or functions between objects of given types (function types).

Declarations

A program in SML consists of a sequence of declarations. Value bindings are introduced by the word `val`. Thus

```
val x = 2;
```

binds the identifier `x` to an object of type: `int`. Function declarations involving variables are introduced by the word `fun` as in

```
fun successor x = x + 1;
```

binding the identifier `successor` to a function of type: `int -> int`. Functions may also be defined with pairs or n-tuples as arguments. Thus

```
fun mult(x,y) = x * y: int;
```

defines a function of type: `(int * int) -> int`.[14] Functions may also be defined in "curried" form as in

```
fun add x y = x + y: int;
```

which defines a function of type: `int -> (int -> int)`. When it is given two integers `add` returns their sum. When it is given a single integer `add` returns a function from integers to integers. Thus the declaration

```
val successor = add 1;
```

is an equivalent way of defining the successor function. Functions may also be defined explicitly in the form

```
val successor = fn x => x + 1;
```

In this form the curried version of integer addition can be written

```
val add = fn x => fn y => x + y: int;
```

[14] The compiler needs to be told whether multiplication is for integers or reals since they have different types.

Lists

A list in SML is an ordered sequence of objects of a given type. Thus `[5,4,7]` is an object of type: `int list`. Lists are constructed from the empty list, denoted by `[]` or `nil`, by inserting a new element at the head. The double colon `::` is used to denote this operation. Thus

```
[5,4,7] = 5 :: [4,7]
        = 5 :: (4 :: [7])
        = 5 :: (4 :: (7 :: nil))).
```

Since every list either matches the pattern `nil` or the pattern `a::l`, where `a` denotes the first element of the list, functions on lists may be defined by structural induction. Thus the sum of a list of integers is defined recursively by

```
fun sum nil = 0
  | sum (a::l) = a + sum l;
```

giving a function of type: `int list -> int`. The more general form of this pattern of declaration is given by the function `iter` defined by

```
fun iter f u nil = u
  | iter f u (a::l) = f a (iter f u l);
```

Replacing `f` by `add` and `u` by `0`, the declaration

```
val sum = iter add 0;
```

defines the same sum function on integer lists. (The function `iter` is more often called `foldr` or `reduce`.)

Two general list handling functions definable in this way, namely `map` and `filter`, are worth mentioning. If `l` denotes a list $[a_1, \ldots, a_n]$ of objects of type: `'a` and `f` is bound to a function f of type: `'a -> 'b`, then the value of `map f l` is the corresponding list $[f(a_1), \ldots, f(a_n)]$ of objects of type: `'b`. (The symbols `'a` and `'b` are used as type variables.) Thus `map` is a function of type: `('a -> 'b) -> ('a list -> 'b list)`. Again if `p` denotes a property of objects of type: `'a`, then `filter p l` is the sublist of `l` of all elements possessing the property in question. Hence `filter` is a function of type: `('a -> bool) -> ('a list -> 'a list)`.

The composition `g o f` of two functions has the obvious meaning. The operator `o` has type: `('b -> 'c) * ('a -> 'b) -> 'a -> 'c`.

We note lastly that a file of SML code could be thought of as written on a single line. All extra spaces, tabs, newlines and indents are ignored. Layout is therefore a matter of style or convenience.

The Code

```
(*******************************************************
 *      Title:      Moebius                            *
 *      LastEdit:   1 June 1987                        *
 *      Author:     Peter M Williams                   *
 *                  University of Sussex               *
 *******************************************************)


datatype SENSE = Inf | Sup;

type LATTICE  = bool list list list;

type DATUM =
     (bool list * (bool list list * bool list list)) * real;


exception hd;
fun hd nil = raise hd
  | hd (a::l) = a;

fun cons a l = a::l;

fun iter f u nil = u
  | iter f u (a::l) = f a (iter f u l);

fun append l m = iter cons m l;

val flat = iter append nil;

fun map f = iter (cons o f) nil;

fun filter p =
    iter (fn a => fn l => if p a then a::l else l) nil;

val sum'r = iter (fn x => fn y => x + y) 0.0;

val inf'r =
    iter (fn x => fn y => if x < y then x else y) (1.0/0.0);
```

```
infix C;
fun (a::l) C (b::m) = (b orelse not a) andalso l C m
  | _ C _ = true;

infix U;
fun (a::l) U (b::m) = (a orelse b) :: (l U m)
  | _ U _ = nil;

val union = iter (fn l => fn m => l U m);

val integrate =
let fun f (a,x) l = if x = 0.0 then (a,x)::l else
    (a,x)::map (fn(b,y) => if b C a then (b,x+y) else (b,y)) l
in iter f nil end;

val invert = map (fn(a,x) => (a,1.0 - x));

val reg =
let fun f l m =
    if m = nil then map (fn(a,w) => (a,exp w)) l else
    let fun g (b,z) n = if z = 0.0 then n else
        map(fn(a,x,y)=> if b C a then (a,x,y+z) else (a,x,y)) n
        val r = iter g (map (fn(a,w) => (a,exp w,0.0)) l) m
        val s = map (fn(a,x,y) => x/y) r
        val k = inf'r s
        val u = map (fn(a,x,y) => (a,x - k*y)) r
        val v = map (fn(a,y) => (a,k*y)) m
    in append u v end
in
integrate o (iter f nil)
end;

fun regularise Inf = reg
  | regularise Sup = invert o reg;

fun abs x = if x < 0.0 then ~x else x;

fun sgn x =
    if x < 0.0 then ~1.0 else if x > 0.0 then 1.0 else 0.0;

fun length'r nil = 0.0
  | length'r (a::l) = 1.0 + length'r l;

exception mean;
fun mean nil = raise mean
```

```
  | mean l = sum'r l/length'r l;

fun center nil = nil
  | center l =
    let val m = mean(map (fn(a,x) => x) l)
    in map (fn(a,x) => (a,x - m)) l end;

fun lookup (a:bool list) nil = 0.0
  | lookup a ((b,x)::l) = if a = b then x else lookup a l;

fun combine f (a::l) (b::m) = f a b :: combine f l m
  | combine f _ _ = nil;

val zero = (map o map) (fn a => (a,0.0));

val add =
    (combine o combine) (fn(a,x) => fn(_,y) => (a,x+y:real));

fun mult k = (map o map) (fn(a,x) => (a,k*x:real));

fun profile sense lattice =
let fun insert (datum as ((b,(pos,neg)),s)) =
    let val x = sgn(s) * ~(ln(1.0 - abs s))
        val w = if sense = Sup then ~x else x
        val (S,T) =
        if sense = Sup then (neg,pos) else (pos,neg)
        val unit  = (hd o hd o rev) lattice
        val c = union unit S
        val l = map (filter (fn a => (c C a))) lattice
        val m =
        iter (fn t => map (filter (fn a => not(t C a)))) l T
        val n =
        (map o map)(fn a => if b C a then (a,w) else (a,0.0)) m
        val q = (flat o map center) n
        fun f(a) = let val ac = a U c in (a,lookup ac q) end
    in (map o map) f lattice end
in
iter (add o insert) (zero lattice)
end;
```

```
abstype MEASURE = Measure of SENSE *
      ((bool list * real) list list * (bool list * real) list)
with
local

fun construct sense (lattice: LATTICE) (data: DATUM list) =
let val profile = profile sense lattice data
    val measure = regularise sense profile
in Measure(sense,(profile,measure)) end

in

val infcon = construct Inf
val supcon = construct Sup

exception sense
infix ++
fun (Measure(s1,(q1,p1))) ++ (Measure(s2,(q2,p2))) =
if s1 <> s2 then raise sense else
let val s = s1
    val q = add q1 q2
in Measure(s,(q, regularise s q)) end

infix **
fun (Measure(s,(q,p))) ** k =
let val kq = mult k q
in Measure(s,(kq, regularise s kq)) end

fun find(Measure(s,(q,p))) = p

end
end;

(************************************************************
The exported functions have types:

val infcon = fn : LATTICE -> DATUM list -> MEASURE
val supcon = fn : LATTICE -> DATUM list -> MEASURE
val ++ = fn : MEASURE * MEASURE -> MEASURE
val ** = fn : MEASURE * real -> MEASURE
val find = fn : MEASURE -> (bool list * real) list
************************************************************)
```

Acknowledgement

I am grateful to Winifred Castle and to colleagues Matthew Hennessy, Mark Millington, Andrew Pitts, Yoshindo Suzuki and Gavin Wraith for advice and encouragement during the preparation of this paper. Any errors or omissions are of course my responsibility.

References

[1] A. Aspect, J. Dalibard and G. Roger. 'Experimental test of Bell's inequalities using time-varying analyzers'. *Physical Review Letters* **49**, 1804–1807, 1982.

[2] J. S. Bell. 'On the Einstein Podolsky Rosen Paradox'. *Physics* **1**, 195–200, 1965.

[3] W. M. Castle, P. M. Williams, P. R. Grob, R. Ellis, N. D. P. Robinson and L. A. Battersby. 'ADRIAN: Adverse Drug Reaction Interactive Advice Network', forthcoming, 1988.

[4] G. Choquet. 'Theory of Capacities'. *Annales de l'Institut Fourier* **5**, 131–295, 1953.

[5] C. Greene. 'On the Möbius algebra of a partially ordered set'. *Advances in Mathematics* **10**, 177–187, 1973; reprinted in G-C. Rota. *Finite Operator Calculus*. Academic Press, 1975.

[6] R. Harper, D. MacQueen and R. Milner. *Standard ML*. Report ECS-LFCS-86-2. Department of Computer Science, University of Edinburgh, 1986.

[7] J. E. Honeycutt. 'On an abstract Stieltjes measure'. *Annales de l'Institut Fourier* **21**, 143-154, 1971.

[8] P. T. Johnstone. *Stone Spaces*. Cambridge University Press, 1982.

[9] A. Joyal and M. Tierney. 'An extension of the Galois theory of Grothendieck'. *Memoirs of the American Mathematical Society'* **309**, 1984.

[10] A. Revuz. 'Fonctions croissantes et mesures sur les espaces topologique ordonées'. *Annales de l'Institut Fourier* **6**, 187–268, 1955.

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

[12] G. Shafer. 'Dempster's rule of combination', unpublished, 1978.

[13] G. Shafer. 'Allocations of probability'. *Annals of Probability* **7**, 827–839, 1979.

[14] G. Shafer. 'Belief functions and possibility measures' in J. C. Bizdek (ed.) *The Analysis of Fuzzy Information* **1**, CRC Press, 1987.

[15] L. Solomon. 'The Burnside algebra of a finite group'. *Journal of Combinatorial Theory* **2**, 603–615, 1967.

[16] Å. Wikström. *Functional Programming using ML*. Prentice-Hall, 1987.