TRUTH OR CONSEQUENCES

Essays in Honor of Nuel Belnap

edited by

J. MICHAEL DUNN
Department of Philosophy and Department of Computer Science, Indiana University, Bloomington, U.S.A.

and

ANIL GUPTA
Department of Philosophy, Indiana University, Bloomington, U.S.A.

KLUWER ACADEMIC PUBLISHERS
DORDRECHT / BOSTON / LONDON

Library of Congress Cataloging-in-Publication Data

Truth or consequences : essays in honor of Nuel Belnap / edited by J. Michael Dunn and Anil Gupta.
p. cm.
Includes bibliographical references and indexes.
ISBN 0-7923-0920-0 (acid-free paper)
1. Logic. 2. Truth. 3. Mathematics--Philosophy. 4. Probabilities. I. Dunn, J. Michael, 1941- . II. Gupta, Anil, 1949- . III. Belnap, Nuel D., 1930-
BC71.T78 1990
160--dc20 90-5141

ISBN 0-7923-0920-0

Published by Kluwer Academic Publishers,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Kluwer Academic Publishers incorporates
the publishing programmes of
D. Reidel, Martinus Nijhoff, Dr W. Junk and MTP Press.

Sold and distributed in the U.S.A. and Canada
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers Group,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

Printed on acid-free paper

Printed in the Netherlands

BAS C. VAN FRAASSEN

FIGURES IN A PROBABILITY LANDSCAPE[1]

1. THE PRIMARY LOGIC OF JUDGEMENT

In general, our opinion is certainly not expressible in precise numerical probabilities. But in some cases it is, and in the case of a small field of propositions — e.g. a field generated by a single proposition — my opinion may be exactly represented by a single probability function. Therefore I shall begin with the fiction that it is always so.

Expression of a judgement is a partial expression of one's opinion. Examples are accordingly:

(a) It seems more likely than not that A
(b) It seems twice as likely that A than that B
(b′) It seems at least as likely that A as that B
(c) It seems π times more likely that A than not, on the supposition that C

The semantic notion needed is that of a probability function which satisfies a given judgement. It is easy to see what that means for this example. For example, p satisfies (a) exactly if $p(A) > p(-A)$.

Epistemic logic has been misguided when it investigated the relation

John believes that . . . $|\vdash$ John believes that ____
In John's opinion, (b) $|\vdash$ In John's opinion, (b′)

which holds if the situation which satisfies the former must also satisfy the latter (*truth preservation*). This led to triviality, because for any example of this relationship we can imagine a moron who is a counter- example. Instead we should look for the significant relationship

(b)$||\vdash$ (b′)

which should hold intuitively if any rational opinion partially expressed through (b) must also be expressible in part by (b′). On our fiction, this means that any probability function which satisfies the former also satisfies the latter.

[1] I am very glad to have this opportunity to honor my teacher, and eventual colleague and friend, Nuel Belnap. In my first year as a graduate student I had Nuel's seminar on the logic of questions, which took us on an exploration of the riches of formal semantics — greater than the fabled treasures of the Indies and the Spanish Main to my eyes. Questions, it turned out, needed a great deal for their understanding: semantics, the technique of consistency and completeness proofs, modal logic, and most of all, the enlargement of philosophical logic beyond the realm of declarative factual statements. That was the beginning; tautological entailment, relevance, algebraic techniques, and much more were to follow. It is easy and pleasant to recall those days, and perhaps most of all Nuel's gentle and unpolemical spirit, always ready to laugh a little at his own and our shared excitement.

J. M. Dunn and A. Gupta (eds.), Truth or Consequences, 345–356.

Brian Ellis (1979) investigated this subject for non-probabilistic judgements; see my (1980).

This sets the primary subject for investigation. We can immediately note a useful reduction of the forms of judgement, in terms of the statistical notion of expectation ("expectation value").

A *space* is a couple $S = \langle K, F \rangle$ with $\Lambda \neq K \in F$ and F a field of subsets of K (the *events* or *propositions*). The probability functions on S are defined by the conditions

(P1) $0 = p(\Lambda) \leq p(A) \leq p(K) = 1$

(P2) $p(A \cup B) + p(A \cap B) = p(A) + p(B)$

A *random variable* (rv) on K is a function $g : K \rightarrow R$ (real numbers) which is "measurable" with respect to F, that is

(RV) $g^{-1}(E) \in F$ for all Borel sets E

where the Borel sets are the countable unions and intersections of intervals ((half-) open, (half-) closed) of real numbers. I shall restrict the discussion to rv with finite range which I shall call simple rv. Then (RV) amounts to

(RV simp) $g^{-1}(r) \in F$ for all real numbers r.

Intuitively, g is a quantity which has a numerical value in each possible state of affairs, and its value can be described by means of field F of propositions.

For given probability function p, the *expectation* $Ep(g) = \Sigma p(A) g_A$ where A ranges over the *characteristic partition* X_g of rv g, which has the members required by (RV simp), and I use the notation

for $A = g^{-1}(r)$, write $g_A = r$.

Example. g measures the daily rainfall in inches, and K is a finite set of days. Then for a person with subjective probability p, the number $E_p(g)$ is the *expected* number of inches of rain — which means, the possible daily amounts of rain, averaged in terms of his probabilities for those amounts.

Let our standard forms of judgement now be all Boolean combinations of

$E(f \geq a)$ satisfied by p iff $E_p(f) \geq a$
$E(f \leq a)$ satisfied by p iff $E_p(f) \leq a$

This is already redundant, since $E(f \leq a)$ is the same as $E(-f \geq -a)$. We can further abbreviate:

$$\begin{aligned} E(f = a) &= E(f \geq a) \,\&\, E(f \leq a) \\ E(f, [a, b]) &= E(f \geq a) \,\&\, E(f \leq b) \\ E(f > a) &= {}^{\sim}E(f \leq a) \\ E(f, (a, b]) &= E(f > a) \,\&\, E(f \leq b) \end{aligned}$$

and so forth. But we can also reduce the other forms of judgement, by using the *indicators* of propositions: $I_A(x) = 1$ if x is in A (A is true at x) and $= 0$ otherwise. Then

$$p(A) = r \text{ iff } E_p(I_A) = r$$

$$p(A) : p(B) = r \text{ iff } E_p(I_A - rI_B) = 0$$
$$p(A|B) = r \text{ iff } E_p(I_{A\cap B} - rI_B) = 0 \qquad \text{provided } p(B) \neq 0$$

because expectation is *linear*:

$$E_p(af + bg) = aE_p(f) + bE_p(g)$$

We can now rewrite our original examples as:

(a) $P[A < -A] = E(I_A - I_{-A} < 0)$
(b) $P[A = rB] = E(I_A - rI_B = 0)$
(b′) $P[A \geq B] = E(I_A - I_B \geq 0)$
(c) $P[A = \pi - A|C] = E\big((I_{A\cap C} - (\pi/\pi + 1)I_C) = 0\big)$

so that the now defined family of judgements comprises a large variety. There are however intuitively possible judgements which are not so expressible, such as that A and B seem independent (which is satisfied by p exactly if $p(A)p(B) = p(A \cap B)$).

2. STATES OF OPINION: VAGUE PROBABILITY

If a person has numerically precise probabilities, the judgements he or she expresses will convey only part of that. The difference between that fiction and ourselves, I propose as improved hypothesis, is that in our case a finite and even small number of judgements may convey all there is to our opinion. But then there is a large class of probability functions which satisfy just those judgements, hence which are compatible with the person's state of opinion. Call that his or her *representor* (*class*).

Suppose my entire state of opinion can be expressed by means of the judgement $P[A \geq 0.5]$ or equivalently $E(I_A, [0.5, 1])$. Then we may equally say *either* that my probability for A is *vague*, with lower and upper bounds 0.5 and 1, *or* that I am *ambivalent* between or about the probability functions p such that $0.5 \leq p(A) \leq 1$. This modelling of vagueness as ambivalence — the "supervaluation" way — is familiar from the general literature on that subject (Fine (1975), Kamp (1977), van Fraassen (1968, 1970) are among the earliest discussions). It certainly has its limits, which have to do with the well-known "vagueness of vagueness", and this subject continues to be explored (see e.g. Tappenden (1989)).

The logic and semantics of vague probability was worked out satisfactorily over several decades (see especially Smith (1961), Levi (1974), Spielman (1976), Williams (1976), Suppes and Zanotti (1977), Jeffrey (1983)). In my opinion the results are perfectly summed up in Theorem 1 of Gaifman (1988). I will explain how his models work, and then develop the generalized theory applying to vague expectation judgements.

An ordinary statement, such as "It rains" is meant to represent some member of the field of propositions on which our probabilities are defined. Each such proposition is represented by a class of situations in which it is true. What about the judgement that it seems as likely as not that it will rain? Its semantic value is the set of probability functions which give 0.5 to that proposition. But we may also wish to consider the proposition *that it is* (*was, will be*) *my opinion that rain seems as likely* *as not*. To have this proposition to think about, it needs to belong to that field. That in turn means that we must think of each situation as including me having some state of opinion or other. Well, that is quite possible. We can think that way without circularity. Here is, roughly, how Gaifman handles it:

> There exists for the space $\langle K, F\rangle$ a function P such that for each A in F and interval $[a, b] \subseteq [0, 1]$, there is a member B of F such that $P(A, [a, b]) = B$, and this function P is moreover such that ...

I will not go into the details, for Gaifman then gives the representation theorem which establishes that this is equivalent to:

> There exists for the space $\langle K, F\rangle$ a function p mapping K into the probability functions defined on F, and such that $P(A, [a, b]) = \{x \in K : a \leq p(x)(A) \leq b\}$ belongs to F for all A in F, and for all a, b in $[0, 1]$.

How exactly does such a structure model my opinion? *First* of all we may think of F as having a subfield for each definite "ordinary" topic on which I have opinions — for instance, the field of propositions about the weather tomorrow, or over the next year, or the tosses of a given die or dice. *Secondly*, we can see my representor class in this structure: the set $RC = \{p(x) : x \in K\}$ is the set of probability assignments (to the whole of F) compatible with my opinion.

There is a problem which I must discuss here, if only briefly. The model represents also opinions about my own opinion. There must be criteria of rationality for those too. Suppose that my opinion about whether it will rain is vague, and completely expressed by the judgement

(1) $P[\mathit{rain}, [0, 0.5]]$

What opinion might I have concerning the autobiographical proposition which is true exactly if (1) expresses my opinion about rain completely? Could I give probability 0.8 to the statement that my probability for rain is greater than 0.5? If I do, my opinion has some defect; the question is only whether it is a factual error or a logical one. Am I like someone who believes that Hitler was a misunderstood good man, or like someone who believes that a square with the same perimeter as a certain circle, has a greater area than that circle? The latter is incoherent, even if he does not realize it. In this paper I shall discuss only opinion "of first order".

If we call such a combination $\langle K, F\rangle$ and P (or p) a model, the general logic of judgements has found an image in a restricted logic of propositions. For example

$$P[A = x], P[B = y], P[A \cap B = z] \;|| \!\!\vdash P[AB = x + y - z]$$

is called correct *because*

$$P(A, [x]) \cap P(B, [y]) \cap P(A \cap B, [z]) \subseteq P(A \cup B, [x + y - z])$$

in *all* models, and hence also, if the left hand side is K in such a model, then so is the right hand side.

3. STATES OF OPINION: VAGUE EXPECTATION

The outline of a theory of vague expectation, along the above lines, is now clear enough. The main first desideratum will be a representation theorem, formulated in such a way that we have a clear and complete axiomatization of the theory. To this end I define the two notions of VEX ("Vague Expectation") model and structure. I will use Greek letters to range over closed real number intervals.

$\langle K, F, p\rangle$ is a *VEX model* iff $\Lambda \neq K \in F$;
F is a field of subsets of K; and
$p : K \rightarrow$ {probability functions with domain F} such that

$$E(f, \delta) = \{x \in K : E_{p(x)}(f) \in \delta\}$$

is in F for each closed interval δ and every simple rv f of the space $\langle K, F\rangle$.

$\langle K, F, E\rangle$ is a *VEX structure* iff $\Lambda \neq K \in F$;
F is a field of subsets of K, and
E : {simple rv} $\times$ {closed intervals} $\rightarrow F$ such that

Ive. $E(f, [inf(f), sup(f)]) = K$
$E(f, \Lambda) = \Lambda$
IIve. $E(kf, [a, b]) = E(f, [ka, kb])$ if $k \neq 0$
IIIve. $\cap E(fi, \delta i) \subseteq E(\Sigma f_i, C)$
where C is the least closed interval that contains all δ_i
IVve. $E(f, [a, b]) \cap E(h, [c, d]) \subseteq E(g, [min(a, c), max(b, d)])$
if $f \leq g \leq h$.
Vve. $E(f, \delta \cap \delta') = E(f, \delta) \cap E(f, \delta')$
VIve. $E(f, \delta \cup \delta') = E(f, \delta) \cup E(f, \delta')$
if $\delta \cup \delta'$ is a closed interval.

Correction to IIIve.

(Read "(i)" as subscript letter "i" and "delta" as the Greek letter delta, in what follows:)

IIIve should end with: "contains all sums of elements a(i) of intervals delta(i)."

In IIIve. I have left the index set indefinite; this should be finite in the present context, but could be countable if we require all the probability functions in question to be probability measures (i.e. countably additive).

We should note that all the other forms of judgement discussed are available to us here. Since a simple rv f has a minimum f_-, for example, $E(f \leq b) = E(f, [f_-, b])$, and of course $E(f > b) = K - E(f \leq b)$, and so forth.

THEOREM 1. *If* $\langle K, F, p\rangle$ *is a VEX model, and* E *is defined by* $E(f, \delta) = \{x \in K : E_{p(x)}(f) \in \delta\}$ *then* $\langle K, F, E\rangle$ *is a VEX structure.*

The proof is elementary, by inspection of the "axioms" Ive.-VIve.

THEOREM 2. *If $\langle K, F, E\rangle$ is a VEX structure, then for each simple rv f of space $\langle K, F\rangle$ and for each set $\Lambda \neq Y \subseteq K$ there are numbers a_f^Y, b_f^Y such that:*

$$Y \subseteq E(f,\delta) \qquad \textit{if and only if} \qquad [a_f^Y, b_f^Y] \subseteq \delta$$

COROLLARY. *If $\langle K, F, E\rangle$ is a VEX structure, then there is for each x in K and each simple rv f of space $\langle K, F\rangle$ a unique number r_f^x such that:*

$$x \in E(f,\delta) \qquad \textit{if and only if} \qquad r_f^x \in \delta.$$

THEOREM 3. *If $\langle K, F, E\rangle$ is a VEX structure and p, p' are defined by*

$$p(x)(A) = inf\{z : x \in E(I_A, [0, z])\}$$
$$p'(x)(A) = sup\{z : x \in E(I_A, [z, 0])\}$$

then $p = p'$ and $\langle K, F, p\rangle$ is a VEX model.

It is easy to see how Theorem 3 will follow from the preceding ones. Applying the Corollary to $f = I_A$ we see at once that $p = p'$ because both pick out r_f^x for each x in K. The "axioms" in the definition of VEX structures give the ordinary characteristics of expectation when the intervals are degenerate; using E' for neutrality:

$$\begin{aligned} E'(f \leq a) &= \{x \in K : E_{p(x)}(f) \leq a\} \\ &= \{x \in K : r_f^x \leq a\} \\ &= \{x \in K : [r_f^x] \subseteq [a_f^K, a]\} \\ &= \{x \in K : \{x\} \subseteq E(f, [a_f^K, a])\} \\ &= \{x \in K : x \in E(f, [a_f^K, a])\} \\ &= E(f \leq a) \end{aligned}$$

and so this set will be in F as required.

So this ends the proof of Theorem 3 from the Corollary. The latter follows from Theorem 2 by setting $Y = \{x\}$. Writing then a_f^x, b_f^x accordingly, we consider the equation

$$E(f, [a_f^x, b_f^x]) = E(f, [a_f^x, r]) \cup E(f, [r, b_f^x])$$

which is true for $a_f^x \leq r \leq b_f^x$. But then, if this interval is not degenerate, x must lie in one of those two parts, and $[a_f^x, b_f^x]$ is not minimal in the required sense — contrary to supposition.

There remains Theorem 2. It will suffice that for $\Lambda \neq X \subseteq K$ there is a smallest closed interval δ such that $X \subseteq E(f,\delta)$. It is already part of the supposition that we have a VEX structure, that this set $E(f,\delta)$ is in the field of propositions.

The rv f is simple, so its range falls inside a closed interval $[f_-, f_+]$ and $E(f, [f_-, f_+]) = K$ by Ive. Consider the family of closed subintervals δ of this interval such that, for given $X \neq \Lambda$, $X \subseteq E(f, \delta)$. By Ive, IVve, Vve this is a proper filter, and since $[f_-, f_+]$ is compact, it follows that this filter has a non-empty intersection. (Cf. e.g. Gaal (1964), Ch. III, sections 1 and 2.) Call it E. This must itself be a closed interval, for if δ is in the family, and $E \subseteq \delta$ then also the least closed interval containing E is contained in δ, for all δ in this filter, and hence part of, and therefore identical with E. This ends the proof.

4. CONDITIONALIZING A VEX

The next obvious question to ask is: how does a person's opinion change with time? There are philosophical disputes and also nice general results in this area, which I have discussed elsewhere (1984, 1986, 1987). At least from a *technical* point of view the old rule of Conditionalization

$$\text{(COND) prior } p \text{, evidence } E \rightarrow \text{posterior } p' = p(-|E)$$

plays a central role almost everywhere. So that is what I shall take up here.

Suppose a person's opinion is vague, equivalently, that his or her opinion is ambivalent on a whole set of probability functions. Suppose in addition that he thinks in a certain case that a certain rule such as COND is applicable. Then it would seem that his posterior opinion should be in effect ambivalence on the new set, formed from the original one by applying the rule to each of its members. This leads us to the definition:

> If $M = \langle K, F, p \rangle$ is a VEX and B in F such that $E(I_B = 0) \neq K$, then the conditionalization of M on B is the structure
>
> $$M_B = \langle K^B, F^B, p^B \rangle$$
>
> where
>
> $$K^B = K \cap E(I_B > 0)$$
> $$F^B = \{A \cap E(I_B > 0) : A \in F\}$$
> $$p^B(x)(Q) = p(x)(Q|B)$$
>
> for all x in K^B, and Q in F^B.

Our suppositions entail at once that $K^B \neq \Lambda$ and that F^B is a subfield of F.

Some violence is done here to the range of envisaged possible situations if, for example, these include ones in which Santa Claus exists, and all the latter are ones in which the precise opinion associated assigns *zero* to B. So it looks as if we are confusing the conditionalization of opinion on B — "learning" that B, taking B as one's total new evidence, with "learning" — taking as one's total new evidence — that one's own opinion *really is* such as to give B a positive probability.

But I think this case does not arise if we make up the VEX in the way it should be made up. Suppose $\langle K^0, F^0\rangle$ represents only possible situations logically independent of my present opinion, but that opinion is vague on the set Q of probability functions on F^0. Then we should build our VEX starting with $K = K^0 \times Q$ — in other words, in such a way that the association of probability functions is used purely and solely to represent our present opinion. But $\langle K^0, F^0\rangle$ itself could be made up of VEXs, say ones that represent in part our possible *future* states of opinion — that makes for no difficulty.

THEOREM 4. *A conditionalization of a VEX is a VEX.*

I chose a VEX model to focus on, because there we see at once that p^B does indeed assign probability functions on F^B. The rest is not so obvious, because there is no definition of conditional probability in terms of expectation. To prove the theorem, I shall first prove a lemma for all VEX models, which concerns a surrogate for conditionalization. Let us define for $M = \langle K, F, p\rangle$:

$$E/^B_x(f) = \sum\{p(x)(A|B)f_A : A \in X_f\}$$

if $x \in E(I_B > 0)$ and undefined otherwise;

and then prove the

LEMMA. *if $x \in E(I_B > 0)$ then $E/^B_x(f) \in [a, b]$ if and only if x belongs to the two propositions $E(h - aI_B \geq 0)$ and $E(h - bI_B \leq 0)$.*

To prove this assume x is indeed in $E(I_B > 0)$ in which case the following are equivalent:

1. $E/^B_x(f) \in [a, b]$
2. $a \leq \sum\{p(x)(A|B)f_A : A \in X_f\} \leq b$
3. $a \leq 1/p(x)(B) \sum\{p(x)(A \cap B)f_A : A \in X_f\} \leq b$
4. $ap(x)(B) \leq E_x(f.I_B) \leq bp(x)B$

because $f.I_B$ defined by $f.I_B(x) = f(x).I_B(x)$ takes value f_A on $A \cap B$ and *zero* elsewhere for exactly the members A of X_f. But 4 is equivalent to the conjunction:

5. $E_x(aI_B) \leq E_x(f.I_B)$ and $E_x(f.I_B) \leq E_x(bI_B)$
6. $x \in E(f.I_B - aI_B \geq 0)$ and $x \in E(f.I_B - bI_B \leq 0)$

as the lemma asserts.

But we see now that there is a close relationship between E^B and its surrogate $E/^B$ partially defined on M. If g is an rv of $\langle K^B, F^B\rangle$ define $g+$ to be the rv on $\langle K, F\rangle$ such that

$$g+_A = g_A \text{ for } A \text{ in } X_g \text{ and } g+_{K-K^B} = 0$$

then we have

7. $E/^B_x(g) = E^B_x(g)$ for all x in K^B
8. $\{x \in K^B : E/^B_x(g) \in [a, b]\} = K^B \cap \{x \in K :$ $E/^B_x(g+)$ is defined and in $[a, b]\}$

9. $$= K^B \cap E(g + .I_B - aI_B \geq 0) \cap E(g + .I_B - bI_B \leq 0)$$

which is certainly in F, and hence, in view of the definition, also in F^B. This ends the proof.

5. PRESERVATION OF FIGURES: FINITE DESCRIBABILITY

We are finite beings. Some of us are also small-minded. Anyone who can be simulated on a computer surely is that. But as long as we are finite then, even if we are not small-minded, with respect to any well-defined field of propositions our expressible opinion must be expressible in a finite number of judgements — so in that respect we are just the same.

Within this context therefore the representor of my state of opinion too is a finite intersection of propositions $E(f \geq 0)$ — let us call such a set a *figure*. (Obviously $E(f \geq a) = E(f - a \geq 0)$ etc.; we are not losing generality here.) The *complexity* of a figure is the least number of such propositions of which it is an intersection.

To say that my representor is a figure means that it is a figure in the set of all probability functions on a given space. If a certain VEX $\langle K, F, p\rangle$ then models my state of opinion, that means that $\{p(x) : x \in K\}$, which is my representor, is a figure in the set of probability functions with domain F.

Question (essentially raised earlier by Gilbert Harman — see my (1987)):

— is this property of being a figure preserved under conditionalization?
— if so, does its complexity decrease, or increase and if the latter, by how much?

It is easy to see why this is of interest: having opted for a representation we believe adequate — partly in its observance of human limitations — we shouldn't like it to be one which becomes inadequate if the state of opinion is changed by something so apparently elementary as conditionalization. But the answer to the first part is *yes*, and the answer to the second is that the complexity *increases at most by a little*, and *often decreases*.[2]

Let X be the set of probability functions on a given domain. A *figure* (in X) — a finitely describable subset — is equivalently the intersection of finitely many "half-spaces" defined by bounded expectation values:

$$E(f \geq a) = \{r : E_r(f) \geq a\}$$
$$E(f \leq a) = \{r : E_r(f) \leq a\}$$

where f is a random variable on the domain. I shall use p, q, r as variables ranging over the probability functions on that domain. Abbreviation:

$$|B| = E(I_B = 1) = E(I_B \geq 1) \cap E(I_B \leq 1)$$
$$= \{r : r(B) = 1\}$$

[2] With thanks to John Broome for a helpful discussion.

where I_A is the *indicator* taking value 1 on A and 0 on $-A$.

For any subset S of X define:

$$S_B = \{p(-|B) : p \in S \text{ and } p(B) > 0\}$$

the *conditionalization* of S *on* B *in* X.

THEOREM 5. *If* S *is a figure in* X, *so is* S_B

To prove this, note that p has a ("orthogonal") decomposition in terms of B, if $0 < p(B) < 1$:

$$p = cp^+ + (1-c)p^-$$

where $0 < c \leq 1$, $p^+ = p(-|B)$; $p^- = p(-|K - B)$.

So also $E_p(f) = cE_{p^+}(f) + (1-c)E_{p^-}(f)$.

LEMMA. $(D \cap E)_C = D_C \cap E_C$.

Obvious from the definition. Hence we need only look at a single half space. I'll do it for $E(f \geq a)$.

LEMMA. $E(f \geq a) = E(f - a > 0)$.

So let us take $S = E(g \geq 0)$. Noting that if p^+ exists then $p(B) > 0$, we argue:

$q \in S_B$ iff $q = p^+$ for some $p \in S$

i.e. such that $E_p(g) \geq 0$

i.e. such that $cE_{p^+}(g) + (1-c)E_{p^-}(g) \geq 0$

i.e. such that $cE_{p^+}(g) \geq -(1-c)E_{p^-}(g)$

Hence: $q \in S_B$ iff $q \in |B|$ and there is r in $|K - B|$ and number $0 < c \leq 1$ such that

$$(*) \qquad E_q(g) \geq \frac{(c-1)}{c} E_r(g)$$

We have two cases:

Case 1: for some r in $|-B|$, $E_r(g) > 0$. Then the RHS of (*) has no lower bound (as c goes to 0 it goes to negative infinity), so $S_B = |B|$

Case 2: for all r in $|-B|$, $E_r(g) \leq 0$. In that case, the RHS of (*) ranges from 0 to positive infinity, so then the LHS need only be non-negative, and $S_B = |B| \cap E(g \geq 0)$.

In both cases S_B is a figure, as was to be proved.

Example. Let S be the set of probability functions on a given domain that give A a probability $> .5$. The rv is then $I_A - .5$, and this has positive expectation for some r in $|-B|$. So the above implies that $S_B = |B|$. Indeed — suppose p gives 1 to B; then $p = p^+$ and whatever $p(A)$ is, we can mix p with a p^- so that the resultant is in S, i.e. gives A a probability $> .5$, while also giving B a non-negative probability.

This was enough to show that the property of being a figure is preserved. Let us now look at the complexity. Conditionalization will transform each half space into either $|B|$ or its intersection with $|B|$. In other words for each relevant rv we either keep the information that its expectation is non-negative or else lose it altogether — while of course gaining the information involved in the new certainty of B. The latter corresponds to the intersection of two half-spaces. Therefore the new complexity is *at most two more* than the old, and generally even less than it was.

THEOREM 6. *The complexity of a figure increases by at most two under conditionalization.*

This sums up all the preceding, for it implies preservation of figurehood.[3]

[3] The author thanks the National Science Foundation for research support.

BIBLIOGRAPHY

Achinstein, P. and Hannaway, O. (eds.) (1985): *Observation, Experiment, and Hypothesis in Modern Physical Science*. Cambridge, MA: Bradford, MIT Press.

Earman, J. (ed.) (1983): *Testing Scientific Theories; Minnesota Studies in the Philosophy of Science*, vol. X. Minneapolis: University of Minnesota Press.

Ellis, Brian (1979): *Rational Belief Systems*. Oxford: Blackwell.

Fine, Kit (1975): "Vagueness, Truth and Logic". *Synthese* 30.

Gaal, S.A. (1964): *Point Set Topology*. New York: Academic Press.

Gaifman, H. (1988): "A Theory of Higher Order Probabilities", pp. 191–219 in Skyrms and Harper, vol. I.

Jeffrey, R.C. (1985): "Probability and the Art of Judgement" in Achinstein and Hannaway, pp. 95–126.

Jeffrey, R.C. (1983): "Bayesianism with a Human Face", pp. 133–156 in Earman (1983).

Kamp, Hans (1977): "Two Theories about Adjectives" in Keenan.

Keenan, E.L. (ed.) (1977): *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press.

Levi, I. (1974): "On Indeterminate Probabilities", *Journal of Philosophy* 71, pp. 319–418.

Martin, R. (ed.) (1970): *The Paradox of the Liar*. New Haven: Yale University Press.

Rescher, N. (ed.) (1987): *Scientific Inquiry in Philosophical Perspective*. Lanham, MD: University Press of America.

Skyrms, B. and Harper W. (eds.) (1988): *Causation, Chance, and Credence*, 2 vols. Dordrecht: Kluwer.

Smith, C.A.B. (1961): "Consistency in Statistical Inference and Decision" *Journal of the Royal Statistical Society* B23, pp. 1–25.

Spielman, S. (1976): "Bayesian Inference with Indeterminate Probabilities" pp. 185–196 in Suppe and Asquith (1976), vol. I.

Suppe, F. and Asquith, P. (1976): *PSA*, in two volumes. East Lansing, MI: Philosophy of Science Association 1976–77.

Suppes, P. and Zanotti, M. (1977): "On Using Random Relations to Generate Upper and Lower Probabilities", *Synthese* 36, pp. 427–440.

Tappenden, J. (1989): Doctoral dissertation, Princeton University.

van Fraassen, Bas C. (1968): "Presupposition, Implication, and Self-Reference", *Journal of Philosophy* 65.

__________ (1978): "Rejoinder: On a Kantian Conception of Language", in Martin.

__________ (1980): "Critical Study of Brian Ellis, Rational Belief Systems", Canadian *Journal of Philosophy* 10, pp. 497–511.

__________ (1987): "Symmetries of Personal Probability Kinematics", in Rescher.

Williams, P. (1967): "Indeterminate Probabilities", pp. 229–246 in Przelewski et. al.