WILLIAM L. HARPER

RATIONAL BELIEF CHANGE, POPPER FUNCTIONS AND COUNTERFACTUALS*

ABSTRACT. This paper uses Popper's treatment of probability and an epistemic constraint on probability assignments to conditionals to extend the Bayesian representation of rational belief so that revision of previously accepted evidence is allowed for. Results of this extension include an epistemic semantics for Lewis' theory of counterfactual conditionals and a representation for one kind of conceptual change.

I. ORTHODOX BAYESIAN PROBABILITY

1. *Preliminaries*

In the orthodox Bayesian tradition of Ramsey, De Finetti and Savage rational belief functions are represented by sharp probabilities.[1] This representation has been defended in a number of ways, but the arguments most characteristic of the tradition turn on the role of belief in guiding decisions.[2] Given some fundamental assumptions about preference and some idealizations and conventions about belief functions, the representation falls out of the Bayesian analysis of rational decision making.

One of the idealizations is that the objects for which the belief function is defined are closed under boolean operations. If $P(A)$ and $P(B)$ exist then so do $P(\bar{A})$, $P(AB)$, etc. The domain of a belief function P will be a boolean algebra $\mathscr{E}$ of propositions. It will be convenient to consider propositions as sets of possible worlds so that $\mathscr{E}$ is a field of subsets of the necessary proposition T. Thus, AB and $\bar{A}$ are the ordinary set operations $A \cap B$ and $T-A$. The possible worlds in T correspond to Savage's (Savage [49] pp. 8–12) possible states of the world and the propositions correspond to his events.[3]

The values of the belief function are real numbers in the interval [0, 1] with the convention that full belief in A is represented by $P(A)=1$.

Ramsey and Savage provide axiomatic characterizations of rational preference according to which a rational agent acts as though his decisions

* Serious difficulties with the construction used in Section III, 3 have been discovered by Robert Stalnaker. See note added in proof to the end of the paper.

Synthese **30** (1975) 221–262. *All Rights Reserved*

were guided by maximizing expected utility relative to a utility function U and a probability function P. The utility function is determined up to positive linear transformations. The probability function is determined uniquely and represents the agents belief function.

Suppose that $A_0 \ldots A_{n-1}$ is a partition of the propositions relevant for a decision among acts $a_0 \ldots a_{k-1}$. The acts do not influence the subjective probability of the propositions.[4] Let $U(a_j A_i)$ represent the utility to the agent of doing act a_j when proposition A_i is the case. The expected utility $E(a_j)$ of each act a_j is the sum over the A_i's of $U(a_j A_i) \cdot P(A_i)$.

$$(1) \qquad E(a_j) = \sum_{i<n} U(a_j A_i) \cdot P(A_i)$$

The preference ordering among the acts for a utility maximizing agent conforms to their expected utilities.

There are various ways of construing the objects of the utility function U. Richard Jeffrey would have $U(a_j A_i)$ attach directly to the proposition that the agent performs act a_j when proposition A_i is the case (Jeffrey [1] pp. 63–81). Savage has utility defined primarily for acts construed as functions mapping possible worlds into consequences, and derivatively for consequences (Savage [49] pp. 17–26). There are other variations. The differences are interesting and important for the problem of axiomatizing preference. (A very nice summary of the field together with the most up to date treatment is to be found in Krantz *et al.* [28] pp. 369–422). For our purposes, however, all that needs attention is a kind of invariance with respect to finer partitions that most treatments share.

Suppose, as before, we have partition $A_0 \ldots A_{n-1}$ of T and acts $a_0 \ldots a_{m-1}$. For each A_i let $B_0^i \ldots B_{k_i-1}^i$ be a partition of A_i such that the acts do not influence the subjective probability of the B_j^i's.

$$(2) \qquad \sum_{i<n} U(aA_i) \cdot P(A_i) = \sum_{i<n} \sum_{j<k_i} U(aB_j^i) \cdot P(B_j^i).$$

What I shall call the finer partitions principle is that the agents acts do not conflict with (2).

2. *Bets and Coherence*

If an agent satisfied Savage's or Ramsey's axioms and his utilities were linear with the stakes in decisions between buying and selling bets then his belief function would determine the prices at which he would buy and sell. He would buy a bet for A at positive stake s when the price is less than $P(A)\cdot s$, sell when the price is more and be indifferent when the price equals $P(A)\cdot s$. De Finetti (De Finetti [12] pp. 103–104) showed that the direct assumption of this connection between bets and beliefs requires that P be a probability function if the agent is to avoid having book made against him. Kemeny (Kemeny [26] pp. 268–269) showed that under this bets-belief assumption having P a probability is sufficient to prevent having book made against the agent. These and later variations by many writers have come to be called dutch book arguments for representing rational belief as a probability function.

A belief function is understood to be coherent as a guide to rational decision making just in case making the bets-beliefs assumption for it does not lead to any system of bets on which the agent faces a total net loss on every outcome. Let $0<x<s$. Any act with utility $s-x$ in outcomes where A holds and $-x$ for outcomes in $\bar{A}$ constitutes buying a bet for A at stake s and price x. Similarly any act with utilities $x-s$ in A and x for outcomes in $\bar{A}$ can be regarded as selling the bet at the same price and stake. The bets-beliefs assumption characterizes the obvious way belief ought to guide choices between acts with such utilities. The dutch book arguments show that this assumption is consistent with expected utility maximizing just in case P is a probability function.

In order to award the utilities properly, actual bets require some procedure for deciding if the proposition is true or false. Brian Ellis has investigated betting systems relative to decision procedures where some propositions may remain undecided (Ellis [11] pp. 131–136). A proposition that may remain undecided is called *semidecidable*. When A is not decided bets for A are called off. They are won or lost as A is decided to be true or false. Ellis shows that in such a framework if arbitrary semi-decidable propositions are allowed then strictly coherent betting ratios cannot be probabilities.[5]

As Ellis sees it the classical dutch book arguments only apply to systems of propositions all of which are decidable, and his result shows that they cannot be generalized to cover semi-decidable propositions. As there are many interesting propositions for which no decision procedures exist Ellis considers his result to be a serious objection to the dutch book justification of the probability representation of rational belief.

In the dutch book arguments, we have been presenting, the acts of buying and selling bets have been defined simply in terms of having appropriate utilities for the outcomes. That actual bets require some method of determining what outcome obtains in order to award utilities properly is beside the point. The idea of coherence is that the bets-beliefs assumption can hold safely under all possible combinations of utilities and outcomes. Ellis' objection just doesn't apply. The classical dutch book arguments had nothing to do with decidability and never were restricted to decidable propositions.

The difficulties with actual bets do not undercut the force of the dutch book argument and Bayesians need not assume that the degree of belief $P(A)$ is behavioristically defined as the critical rate a which he would buy and sell actual bets on A. These difficulties do indicate, however, that measuring degrees of belief is not as straight-forward as one might hope.

Ramsey's treatment suggests a way to use gambling devices and prizes to measure degrees of belief (Ramsey [46] pp. 77–79). Suppose that the outcomes of some gambling devices are themselves value neutral to the agent and that he assigns equal subjective probability to the designed equi-probable outcomes. Let c be a prize that the agent desires and consider the choice between (a) and (b).

(a) receive c, if A and nothing if $\bar{A}$

(b) receive c, if gambling device comes up with any one of m out of the n outcomes, and nothing if it comes up with one of the rest.

The agents degree of belief is measured as close as one wants by considering his preferences between (a) and (b) in various choices of this sort. Since Ramsey's time many such procedures have been proposed, some of which are more sophisticated than this (cf. Krantz, *et al* [28] pp. 900–901 for discussion and further references). All of the available procedures, however, share with the one we sketched that sometimes the choices will involve gambles on propositions for which no convenient method of verification exists.

Many interesting propositions have no convenient procedures for verification. Popper points out that most general scientific hypotheses have special problems. According to him, there is no method for finding out that they are true, but there may be a method for finding out that they are false (Popper [43] pp. 27–48, [45] pp. 1–30 and many further publications referenced in [45]). If A cannot be verified then any attempt to build explicit tests for deciding A into the choice between (a) and (b) is doomed. Considerations of this sort have led Abner Shimony to give up the expected utility justification for inductive probability (Shimony [51] pp. 103–104).

I think it is important to see that when A is not conveniently verified the option (a) is considerably different than a bet on a horse race where the outcome will be disclosed without fail. The decision between (a) and (b) is a kind of thought experiment of what one would do under the hypothetical assumption that choosing (a) would without fail result in pay off (c) if A and nothing if $\bar{A}$. I do not, however, see why such hypothetical choices cannot be made in all seriousness. Moreover, the degrees of belief that they reveal are exactly those that would operate to guide decisions where the agent thinks that the truth of A matters.

3. *Acceptance and Strict Coherence*

If P is a classical probability on $\mathscr{E}$ then $\Delta p = \{A : P(A) = 1\}$ has some important characteristics. For all $A, B \in \mathscr{E}$

(1) If $A, B \in \Delta p$ then $A \cap B \in \Delta p$
(2) If $A \in \Delta p$ and $A \subseteq B$ then $B \in \Delta p$
(3) $A \cap \bar{A} = \emptyset \notin \Delta p$.

A subset Δ of $\mathscr{E}$ satisfying these requirements is a proper filter of $\mathscr{E}$ and corresponds to a consistent set of propositions closed under semantical consequence. The semantical consequence relation is that of truth preservation. A proposition A is true at world W just in case $W \in A$. Clearly 1 preserves truth in that whenever both A and B are true in W so is $A \cap B$. Similarly for 2. Constraint 3 is consistency.

We have introduced a convention that $P(A) = 1$ represents full belief in A. Robert Stalnaker considers a belief function P to represent an idealized possible state of knowledge where every A such that $P(A) = 1$ is known by the agent (Stalnaker [52] p. 66). The concept of knowledge is idealized in that what the agent knows is closed under semantical consequence just as in Hintikka's analysis (Hintikka [21] pp. 8–39). When P is so considered the set $K(P) = \cap \Delta p$ consists in exactly those worlds that are epistemically possible relative to P in Hintikka's sense of epistemic possibility (*Ibid.*, and Hintikka [22]). An agent who knows every A in Δp in effect knows that the actual world is in $K(P)$. I interpret $P(A)=1$ as the agent accepts A, and interpret $K(P)$ as a proposition that expresses the total content of what he accepts. Stalnaker's heuristic of a possible state of knowledge is a good one, however, because the agent will usually regard himself as knowing just those propositions he accepts.

Let h be a function mapping $\mathscr{E}$ into $K(P)$ so that for each A in $\mathscr{E}$

$$h(A) = A \cap K(P)$$

and let $\mathscr{E}_p$ be the range of h. $\mathscr{E}_p$ is a field of subsets of $K(P)$ isomorphic to the quotient algebra of $\mathscr{E}$ modulo the filter Δp, and h is a homomorphism of $\mathscr{E}$ onto $\mathscr{E}_p$. Where we have a decision D relative to $\mathscr{E}$ with partition $A_0 \ldots A_{n-1}$ of T and acts $a_0 \ldots a_{k-1}$ let $h(D)$ be the corresponding decision in $\mathscr{E}_p$ with partition $h(A_0) \ldots h(A_{n-1})$ of $K(P)$ and expected utility $E_h(a_j) = \Sigma_{i<n} u(a_j h(A_i)) . P(h(A_i))$ for each act a_j.

Remark 2.1. D and $h(D)$ are equivalent in that for every a_j

$$\underline{E}(a_j) = \underline{E}_h(a_j).$$

Proof: Break $\underline{E}(a_j)$ down into a sum over $h(A_i)$'s and a sum over $(A_i - h(A_i))$'s. The later sum is zero. ■ Every decision relative to partitions of T in field $\mathscr{E}$ is equivalent to its corresponding decision relative to partitions of $K(P)$ in field $\mathscr{E}_p$. The expected utility framework does not distinguish between $U(a_j A_i)$ and $U(a_j A_i \cap K(P))$.

There are two opposed ways of dealing with this point and they represent two quite different approaches to the Bayesian analysis of rational belief. On the one hand, we may say that no rational agent would accept contingent propositions because this would be tantamount to ignoring what could happen in some of the possible outcomes. On the other hand, we may allow that rational agents do accept contingent propositions and hold that for a rational agent only those outcomes consistent with what he knows are relevant for making his decisions. The second position sees the Bayesian belief function as an extension of our ordinary ideas of belief and knowledge. In addition to representing what the agent regards himself as knowing it also assigns degrees of belief to those propositions his knowledge does not decide. On this view the agent may be quite rational to accept any of the propositions we would usually regard him as knowing. The first view is the strict-coherence position. On it the ordinary notions of belief and knowledge are inoperative for decision making. They are replaced by the more adequate notion of rational partial belief (see Richard Jeffrey [24] for an excellent defense of this view). Those cases where an agent would normally be regarded as accepting some contingent propositions are really just cases where his degree of belief is close to but not equal to 1 (see Teller [55] p. 240).

An agent who thinks he rationally accepts some contingent truths might argue:

> "I am sure that my hand is before me on the page, and that the population of the United States is greater than that of Canada."

The strict coherence advocate replies:

> "Then you should be indifferent between paying $1000 for a change to lose it if you are wrong and get it back if you are right, and paying nothing for a chance to get $1000 if you are wrong and nothing if you are right."

If the agent refuses to see the light, the strict coherence advocate simply increases the amount. Sooner or later the agent breaks down and admits that he is just a little bit unsure.

At first glance the case for strict coherence seems strong. But, the situation is less simple than it may seem. Try the Ramsey degree of belief measurement. Consider a choice between

(a) Receive $1000 if my hand is really there on the page and nothing if not.

(b) Receive $1000 if the random device comes up on any but *m* of the *n* possible outcomes and nothing if not.

No matter how small m/n is made, so long as there is at least one unfavorable outcome, I will prefer (a) to (b). This also holds for the proposition that the population of Canada is less than that of the United States.

If the agents degree of belief is 1 why isn't he indifferent between the choices offered by the strict coherence advocate? One answer might be that the utilities of receiving a net of zero dollars can differ with the choice context. If the agent breaks down at one stake, but not at another then one has prima facie evidence that the stake can change the relative desirabilities of zero dollar net gain in different contexts. Even if the choice offered by the strict coherence advocate could avoid all difficulties of this sort a result in his favor would be a puzzle for decision theory rather than an unambiguous defense of his position.

Strict coherence is usually defined and defended relative to the bets-belief assumption. A belief function is strictly coherent just in case the bets-belief assumption does not lead to any system of bets where the agent suffers a net gain on no outcome and a net loss on some outcome. From Kemeny's result we have that a belief function is strictly coherent just in case it is coherent and $P(A)=1$ just for those propositions that are true in every *relevant possible outcome.* A bet having utilities matching the money offered in the strict coherence advocate's choice would lead to a violation of strict coherence if the agent accepts a contingent proposition and the relevant outcomes are all the worlds in T. Stalnaker suggests the obvious way to have acceptance of contingent propositions not violate strict coherence (Stalnaker [54] p. 68). The relevant possible outcomes are just those in $K(P)$.

Strict coherence was originally introduced by Shimony as a constraint on confirmation functions (Shimony [50] pp. 9–12). A confirmation function $\mathscr{C}$ on $\mathscr{E}$ maps $\mathscr{E}\times\mathscr{E}-\{\emptyset\}$ into the reals and satisfies the following basic axioms:

(1) $0 \leqslant \mathscr{C}(H/E) \leqslant 1$

(2) $\mathscr{C}(E/E) = 1$

(3) $\mathscr{C}(H/E) + \mathscr{C}(\bar{H}/E) = 1$

(4) If $E \cap H \neq \emptyset$ then $\mathscr{C}(H \cap J/E) = \mathscr{C}(H/E) \,.\, \mathscr{C}(J/E \cap H)$.

I have used Carnap's axiomatization (Carnap [7] p. 38), because confirmation functions have been so closely identified with Carnap's program of inductive logic. The motivation for the constraints on confirmation is that $\mathscr{C}$ is to be a conditional probability adequate to play the following role:

> If an ideally rational agent were to have exactly E as his total evidence then his degree of belief in proposition H ought to be $\mathscr{C}(H/E)$.

Carnap's program may be regarded as the attempt to put as many constraints on $\mathscr{C}$ as can be justified by this role.[6]

When Shimony applied the strict coherence argument to confirmation functions he made $\mathscr{C}(H/E)$ determine betting rates for bets on H when the relevant possible cases are restricted to just those in E. On this procedure a confirmation function is strictly coherent just in case it is regular. Regularity is the constraint that

$$\mathscr{C}(H/E) = 1 \quad \text{only if} \quad E \subseteq H.$$ [7]

The heuristics of confirmation indicate that regularity is a natural constraint on confirmation functions. Suppose $\mathscr{C}(B/A)=1$ but $A \nsubseteq B$. Thus, proposition $\bar{A} \cup B$ is contingent (i.e. $\bar{A} \cup B \neq T$). For any E and H $\mathscr{C}(H/E)=\mathscr{C}(H/E \cap (\bar{A} \cup B))$. If $E \neq E \cap (\bar{A} \cup B)$ then the heuristic of confirmation is violated since $\mathscr{C}(H/E)$ corresponds to having $E \cap (\bar{A} \cup B)$ as evidence rather than just E.

When $\mathscr{C}$ is regular the corresponding absolute probability function,

$$\mathscr{C}(H) = \mathscr{C}(H/T),$$

is strictly coherent in the strict sense. This does not, however, imply that the belief function of a rational agent who guides his beliefs by $\mathscr{C}$ would also be strictly coherent in the strict sense. Carnap represents rational belief of an agent with background evidence as a credence function. Where K is the total content of the agent's background evidence his rational credence function $\mathscr{C}_{(K)}$ conforms to $\mathscr{C}$ so that

$$\mathscr{C}_{(K)}(H/E) = \mathscr{C}(H/E \cap K)$$

and

$$\mathscr{C}_{(K)}(H) = \mathscr{C}(H/K).$$

Clearly $\mathscr{C}_{(K)}$ need not be strictly coherent in the strict sense.[8] Moreover, $\mathscr{C}_{(K)}$ need not even be regular. If $K \subseteq H$ then $\mathscr{C}_{(K)}(H/E)=1$ even if $E \nsubseteq H$.

This difference between $\mathscr{C}_{(K)}(H/E)$ and $\mathscr{C}(H/E)$ corresponds to the following different heuristics for conditional probability. The confirmation conditional probability $\mathscr{C}(H/E)$ is what the rational agent would assign to H if his total evidence were reduced to nothing but the proposition E. The credence conditional probability $\mathscr{C}_{(K)}(H/E)$ is what the rational agent would assign to H if E and nothing further were added to what he now accepts.

4. *Conditionalization and Learning from Experience*

One of the most striking features of the orthodox Bayesian tradition is the representation of learning from experience by conditionalization. A change from rational belief function P_0 to P_1 is by conditionalization on A just in case $P_1(B)$ is the conditional probability $P_0(B/A)$ for every B. Let us suppose that $P_0(A)>0$ and examine the claim that $P_0(AB)/P_0(A)$ is the appropriate new degree of belief in B for a rational agent who has just altered his belief function P_0 by learning A and nothing more.

The most extensive discussion of such claims is by Paul Teller (Teller [55, 56]). Under assumptions that come to the specification that $P_0(A)>0$ and that accepting A is the total direct epistemic input from the learning experience Teller suggests that if $P_0(B)=P_0(C)$, $B\subseteq A$ and $C\subseteq A$, then it ought to be that $P_1(B)=P_1(C)$ (Teller [55] pp. 233–238). He shows that this qualitative assumption about rational belief change is equivalent to conditionalization under fairly normal structural assumptions about belief functions (Teller [55] pp. 223–230). Teller, also, reports a rather ingenious dutch book argument by David Lewis (Teller [55] pp. 222–225).

I shall give a dutch book argument based on an idea that can be used to help extend the representation of rational belief so that conditionalization on propositions of zero probability is allowed. The idea can also be used to defend Savage's conditional expected utility argument for conditionalization from an objection Teller makes against it (Teller [56] p. 18).

Suppose that $P_0(A)>0$ and P_1 is the appropriate new belief function when the change from P_0 is to learn A and nothing more. The new set of propositions accepted $\Delta(P_1)$ ought to be $\Delta(P)\cup\{A\}$ and the content of what is accepted $K(P_1)$ ought, therefore, to be $K(P_0)\cap A$. The basic idea is that the shift from P_0 to P_1 ought to be as minimal as is required for accepting A. One obvious principle governing minimality here is that one not give up any proposition he already accepts unless he needs to. Since $P_0(A)>0$; $K(P_0)\cap A\neq\emptyset$ and A is compatible with everything the agent accepts. Thus the agent need give up none of the propositions in $\Delta(P)$ when he accepts A.

When the bets-beliefs assumption is applied to P_1 the relevant outcomes are just those in $K(P_1)$. A conditional bet for B on A is one that is called off in relevant outcomes not in A and that is won or lost as usual for relevant outcomes in A. Since $K(P_1)=K(P_0)\cap A$ any bet for B at stake s and price $\gamma\cdot s$ relative to $K(P_1)$ outcomes has exactly the same net return on every outcome as a conditional bet for B on A at the same price and stake with $K(P_0)$ relevant outcomes. If the bets-beliefs assumption holds for both P_0 and P_1 then $P_1(B)$ must be the same as the critical price ratio for conditional bets for B on A relative to $K(P_0)$ outcomes. De Finetti (De Finetti [12] pp. 108–109) has shown that the P_0 critical price ratio for conditional bets for B on A must be $P_0(AB)/P_0(A)$ if the agent is to avoid dutch books.[10]

I think that implicit assumptions like those I make explicitly here account for the fact that some Bayesian writers (e.g. De Finetti) were content to limit their argument for conditionalization to showing that the ratio is the appropriate betting rate for conditional bets.

If all rational learning from experience is by conditionalization on new evidence and belief functions are only classical probabilities, then no way is provided for revising previously accepted evidence on the basis of new inputs. Suppose P_1 arises from P_0 by conditionalization on A. Then, $P_1(A)=1, P_1(A/C)=1$ for every C such that $P_1(C)>0$, and no conditional probability $P_1(A/C)$ exists for any C such that $P_1(C)=0$. Clearly all revision of the assignment $P_1(A)=1$ blocked. Any hypothetical new evidence C that is not already rejected will continue to support A, and any hypothetical C that is rejected cannot play a new evidence role because the relevant conditional probabilities do not exist.

Richard Jeffrey (Jeffrey [23] pp. 153–164) has proposed a generalization of conditionalization according to which $P_1(A)$ need not be 1. A change from P_0 to P_1 originates in the partition $A_0 \ldots A_{n-1}$ just in case for $i<n$ $P_1(B/A_i)=P_0(B/A_i)$ for every B. When this happens the coherence constraints on P_1 generate the rule:

$$P_1(B) = \sum_{i<n} P_0(B/A_i)\cdot P_1(A_i).$$

Jeffrey argues that there can be cases where one rationally responds directly to experience by shifting $P_0(A)$ to some new value $P_1(A)$ without accepting A or anything else as new evidence. He claims that in such cases the partition $\{A, \bar{A}\}$ should be an origin for the shift from P_0 to P_1.

Isaac Levi, a defender of rational acceptance, rightly saw that Jeffrey's rule would provide for a representation of learning from experience where no contingent evidence would ever need to be accepted (Levi [34]). His attack on Jeffrey's rule, however, was defective (Harper and Kyburg [20], Levi [35]). Teller has now provided a quite adequate defense of Jeffrey's rule (Teller [55] pp. 243–257). Levi's fears have beeen realized. The strict-strict coherence approach can handle learning from experience. In fact it handles it better than the ordinary framework of conditionalization and acceptance because any change is open to correction on the basis of future observations.

I think that the standard notions of acceptance and bodies of evidence are too useful to give up. But, if one is to accept contingent evidence of the usual sort some provision must be made for revision of previously accepted evidence. A first step toward this is to allow $P(B/A)$ to be defined even when $P(A)=0$.

II. EXTENDING THE REPRESENTATION OF RATIONAL BELIEF TO POPPER FUNCTIONS

1. *Popper's Probability Functions*

Popper provides an axiomatic treatment of probability in which conditional probability is primitive and exists everywhere.[11] Suppose $\mathbf{F}$ is a minimal algebra with a binary operation AB and unary operation $\bar{A}$.[12] Nothing specific about the algebraic properties of these operations is assumed. The following axioms characterize one version of a Popper probability function P mapping $\mathbf{F}\times\mathbf{F}$ into the reals.[13] For all A, B and C in $\mathbf{F}$,

a1. $0 \leqslant P(B/A) \leqslant P(A/A) = 1$

a2. If $P(A/B) = 1 = P(B/A)$ then $P(C/A) = P(C/B)$

a3. If $P(C/A) \neq 1$ then $P(\bar{B}/A) = 1 - P(B/A)$

a4. $P(AB/C) = P(A/C)\cdot P(B/AC)$

a5. $P(AB/C) \leqslant P(B/C)$.

Popper adds the additional constraint that there be some C and D in $\mathbf{F}$ such that $P(A/B) \neq P(C/D)$. I shall call functions satisfying these requirements Popper functions.

In the classical mathematical treatment probability is defined as a non-negative additive set function normalized to 1. Suppose that $\mathscr{E}$ is a field of subsets of some non-empty set T. A function M mapping $\mathscr{E}$ into the reals is a classical probability just in case[14]

(i) $M(T) = 1$, and
(ii) If $A \cap B = \emptyset$ then $M(A \cup B) = M(A) + M(B)$.

Sometimes the treatment is generalized to have $\mathscr{E}$ an arbitrary boolean algebra. In this case T will be the maximum of $\mathscr{E}$ and in place of set intersection and union we will have boolean meet and join.

Classical conditional probabilities are introduced by definition as ratios of absolute probabilities.

(iii) $M(B/A) = M(AB)/M(A)$,

provided $M(A) > 0$. If $M(A) = 0$ then no classical probability $M(B/A)$ exists. In a Popper function conditional probability is primitive, but absolute probability is easily represented. For Popper function P the absolute probability $P(A)$ is conditional probability relative to $T = \overline{A\bar{A}}$, so that

$$P(A) = P(A/T)$$

for A in $\mathbf{F}$. Where $P(A) > 0$ the Popper conditional probability $P(B/A)$ is a ratio of absolute probabilities.

$$P(B/A) = P(AB)/P(A)$$

just as classical conditional probability. The most salient difference between a Popper function and classical probability is that $P(B/A)$ exists even when $P(A) = 0$.

Popper's extension of conditional probability to all pairs of elements has some mathematical advantages. Chief among these is that P induces an interesting boolean algebra of equivalence classes on the minimal algebra $\mathbf{F}$. For elements A, B of $\mathbf{F}$ define,

d1.3. (i) $A \underset{p}{\sim} B$ iff $P(A/C) = P(B/C)$ for all C in $\mathbf{F}$
(ii) $[A]p = \{C \in \mathbf{F} : A \sim_p C\}$
(iii) $F/P = \{[C]p : C \in F\}$

When $A \underset{p}{\sim} B$ we say that A is P-equivalent to B. The subset $[A]p$ of $\mathbf{F}$ is the equivalence class of A under P, and $\mathbf{F}/P$ is the set of equivalence classes induced by P on **F**. The following operations are defined on $\mathbf{F}/p$

(iv) $[A]p \wedge [B]p = [AB]p$

(v) $\overline{[A]}p = [\bar{A}]p$

to form the algebraic structure $\mathbf{F}//p$. Popper proves the following theorem.

t1.1. $\mathbf{F}//p$ is a boolean algebra with $[A]p \wedge [B]p$ as meet and $\overline{[A]}p$ as complement for A, B in **F**.

This theorem shows that the constraints on P are sufficient to impose boolean behaviour on the unstructured operation of **F**.

In the classical treatment $\mathscr{E}$ must already be a boolean algebra, before M can be defined on it. As Popper points out, the algebraic structure of $\mathscr{E}$ is an additional assumption buried in the classical characterization of probability. The **F** of a Popper function need only be a minimal algebra, and the algebraic properties used in probability reasoning are generated by the explicit constraints on P.

The introduction of $\mathbf{F}//p$ allows us to formulate some further connections between Popper functions and classical probabilities. Let P_A be defined on **F** so that $P_A(B) = P(B/A)$ and $P_{[A]}$ be the corresponding function on $\mathbf{F}//p$ so that $P_{[A]}([B]) = P_A(B)$. We have the following remarks

r1.2. (i) If $P(\bar{A}/A) \neq 1$, then $P_{[A]}$ is a classical probability on $\mathbf{F}//p$
(ii) If $P(\bar{A}/A) = 1$, then P_A (hence $P_{[A]}$) is the incoherent constant function assigning 1 to every element.

The absolute probability $P(A)$ is simply $P_T(A)$ and corresponds to the classical probability $P_{[T]}$ on $\mathbf{F}//T$.

Two elementary properties of the other extreme value for Popper functions are also of interest.

r.1.3. (i) $P(A/\bar{A}) = 1$ iff $P(A/C) = 1$ for all C
(ii) $P(A/\bar{A}) \neq 1$ iff $P(A/C) = 0$ for some C.

The first of these has the effect that the maximum of $\mathbf{F}//p$ is the set of all A such that $P(A/\bar{A}) = 1$. We shall say that A is P-valid just in case $P(A/\bar{A}) = 1$. The second remark is that whenever A is not P-valid there is some C such that $P(A/C) = 0$. This will be useful in showing P-validity by indirect proof.

2. *Stalnaker's Representation Theorem*

Robert Stalnaker has constructed a dutch book representation theorem for Popper functions as coherent extended conditional belief functions (Stalnaker [52] pp. 70–74). The representation is based on his idea that the relevant outcomes for strict coherence depend on what the agent accepts. The present treatment differs from Stalnaker's in formulation and in the details of the proof. The most important difference is the explicit emphasis on the way the constraints on $K(P_A)$ justify axioms 2 and 3 for Popper functions.

Let us extend the representation of rational belief so that $P(B/A)$ is defined for every pair A, B in $\mathscr{E}$. We want $P(B/A)$ to represent the degree of belief that would be rational for the agent to assign to B were he to accept A as his total new input from experience. We shall think of P_A (the function defined on $\mathscr{E}$ so that $P_A(B)=P(B/A)$) as the absolute belief function the rational agent would have were he to minimally revise his beliefs to accept A. In light of this motivation certain general constraints on rational extended conditional belief functions seem warranted.

Let

$$\Delta(P_A) = \{B: P(B/A) = 1\}$$

and

$$K(P_A) = \bigcap \Delta(P_A).$$

Just as with classical conditionalization $\Delta(P_A)$ is to be the set of propositions the agent would accept if his absolute belief function were P_A and $K(P_A)$ is the set of worlds where all these propositions hold. Since P_A is to be a belief function where A is accepted it is required that $P(A/A)=1$. This gives the constraint

(1) $\quad K(P_A) \subseteq A.$

We shall express the constraints in terms of $K(P_A)$ where possible.

Another obvious constraint is that $K(P_A)$ be nonempty for at least some A.

(2) (i) $\quad K(P_A) \neq \emptyset, \quad$ for some $\quad A.$

If $K(P_A)$ is empty then A is regarded as absurd in that there is no world consistent with all the propositions the agent would be committed to were he to minimally revise his beliefs to accept A. It is convenient to have a convention for cases where $K(P_A)$ is empty.

(2) (ii) If $K(P_A) = \emptyset$, then $P(B/A) = 1$ for all B.

This convention corresponds to the idea that anything follows from something that commits you to a contradiction.

The main constraint is that P_A be coherent when the possible cases are restricted to $K(P_A)$.

(3) P_A is coherent relative to $K(P_A)$ if $K(P_A) \neq 0$.

The justification for this is the obvious one that P_A is to be a belief function appropriate to guide the agent's decisions relative to partitions of $K(P_A)$.

Our motivation that P_A be a minimal revision to accept A and the classical conditionalization property that $K(P_A) = K(P) \cap A$ when $P(A) > 0$ suggest a further constraint.

(4) $K(P_{AB}) = K(P_B) \cap A$, provided $K(P_B) \cap A \neq \emptyset$.

If $K(P_B) \cap A$ is non-empty then a minimal revision of $K(P_B)$ to accept A is simply to add A to what one already accepts.

The justification of this last constraint corresponds to the principle applying to minimal revisions of belief functions. When revising your beliefs in order to accept A do not give up anything you already accept unless you need to.

We shall take any function from $\mathscr{E} \times \mathscr{E}$ into the interval [0, 1] which satisfies 1–4 to be a suitable representation for an extended conditional belief function (ebf).

THEOREM 2.1: If P is an extended condition belief function then P is a Popper function.

Proof: The plan of the proof is to show that a violation of any of the axioms for Popper functions will also violate one of the constraints on extended conditional belief functions. Usually this will consist in showing a violation of coherence by means of one of the betting systems used in John Kemeny's version of the dutch book argument (Kemeny [26] pp. 263–266).

Axiom 1 for Popper functions,

a1. $0 < P(B/A) < P(A/A) = 1$,

is trivially required by constraint 1 on extended conditional belief functions.

Consider axiom 3.

a3. If $P(C/A) \neq 1$ then $P(\bar{B}/A) = 1 - P(B/A)$.

When the hypothesis is satisfied for some C, constraint (2) (ii) on conditional belief functions requires that $K(P_A)$ be non-empty. Therefore, constraint 4 non-trivially requires coherence relative to $K(P_A)$. Unless $P(\bar{B}/A) = 1 - P(B/A)$ coherence relative to $K(P_A)$ will be violated.

The situation with axiom 5

a5. $P(AB/C) \leqslant P(B/C)$

is similar. If $K(P_C) = \emptyset$ then both sides of the inequality are trivially 1. If $K(P_C) \neq \emptyset$ then coherence with respect to $K(P_C)$ is sufficient to guarantee that the inequality holds. The axioms so far considered all fall out directly from the coherence requirement that P_A be a classical probability function on $\mathscr{E}/P_A$.

Axioms 2 and 4 involve relations between different $K(P_A)$'s. Let us deal with a4 first.

a4. $P(AB/C) = P(A/C) \cdot P(B/AC)$

There are two cases to consider.

Case I. $K(P_C) \cap A = \emptyset$. When $K(P_C) \cap A$ is empty then both sides of a4 must be zero. Since A is false in every world in $K(P_C)$ both $P(A/C)$ and $P(AB/C)$ equal zero.

Case II. $K(P_C) \cap A \neq \emptyset$. When $K(P_C) \cap A$ is non-empty then condition 4 on conditional belief functions requires that $K(P_{AC}) = K(P_C) \cap A$. This has the effect that a conditional bet on B relative to A in $K(P_C)$ at odds $r : 1 - r$ and stake S has exactly the same outcomes as a straight bet on B at the same odds and stake in $K(P_{AC})$. Given this, bets on B in $K(P_{AC})$ can be represented as conditional bets in $K(P_C)$ so that all the degree of belief values can be represented in a system of bets all relative to the same partition of possible outcomes. If a4 is not satisfied the incoherence of P will show up in the bets used by Kemeny to show the corresponding law for confirmation functions (Kemeny [1] pp. 265–266). Though axiom 2 is not as familiar to students of probability as the other axioms it does follow from the constraints on extended conditional belief functions.

a2. If $P(B/A) = 1 = P(A/B)$ then $P(C/A) = P(C/B)$

Assume $P(B/A)=1=P(A/B)$. By condition 4 we have $K(P_{AB}) = K(P_A)\cap B$ if $K(P_A)\cap B\neq\emptyset$, and $K(P_{AB})=K(P_B)\cap A$ if $K(P_B)\cap A\neq\emptyset$. Case I: $K(P_A)\cap B\neq 0$ and $K(P_B)\cap A\neq 0$. Here both $K(P_A)$ and $K(P_B)$ equal $K(P_{AB})$ so that a2 follows easily by coherence constraints on $K(P_{AB})$. Case II: $K(P_A)\cap B=\emptyset$. Here $K(P_A)$ is empty, since by the hypotheses of the theorem $P(B/A)=1$. Therefore, by constraint (2) (iii) $P(\bar{A}/A)=1=P(\bar{A}/\bar{A})$.

The general constraint on Popper functions

r1.3 (i). If $P(D/\bar{D}) = 1$ then $P(D/E) = 1$ for all E

follows from axioms a1 and a3-a5 and constraint 1–4 on ebf's.[15] Since these already established axioms are sufficient for it, we may use r1.3 (i) to justify a2. Using this remark we have $P(\bar{A}/B)=1$, since $P(\bar{A}/\bar{\bar{A}})=1$.

We now have $K(P_B)$ is empty, since both $P(A/B)=1$ and $P(\bar{A}/B)=1$. Thus, $K(P_B)=K(P_A)$ and both are empty. Case III $K(P_B)\, A=\emptyset$ is symmetrical with what we showed for case II. This completes showing the validity of a2 and completes the proof of the theorem.■

We say that Popper function P on field $\mathscr{E}$ is *compact* just in case for all A

$$K(P_A) = \emptyset \quad \text{only if} \quad P(\bar{A}/A) = 1 .$$

If $\mathscr{E}$ allows filters Δ such that $\bigcap\Delta=\emptyset$, but $\emptyset\notin\Delta$ then there can be Popper functions on $\mathscr{E}$ that fail to be compact. Such Popper functions will violate condition 2ii on extended conditional belief functions.

THEOREM 2.2: If P is a compact Popper function on $\mathscr{E}$ then P, is a suitable representation for an extended belief function on $\mathscr{E}$.

Proof. Constraints 1–2ii are trivially met. Constraint 3 follows from the fact that P_A is a probability function whenever $K(P_A)\neq\emptyset$ together with Kemeny's result (p. 223 above). Constraint 4 follows by manipulation from the Popper function axioms.[16]■

Stalnaker constructed extended belief functions and Popper functions on sentences. Suppose **S** is a set of sentences closed under syntactical operations **ab** and $\bar{\mathbf{a}}$. Let T be the set of maps v from **S** into $\{0, 1\}$ such that $v(\mathbf{ab})=v(\mathbf{a})\cdot v(\mathbf{b})$ and $v(\bar{\mathbf{a}})=1-v(\mathbf{a})$ for **a**, **b** in **S**. This is the set of truth valuations on **S** with **ab** as conjunction and $\bar{\mathbf{a}}$ as negation. For each **a** let A be the set of v in T such that $v(\mathbf{a})=1$. The set $\mathscr{E}$ of A such that **a** is in **S** is a field of subsets of T.

A function P taking $\mathbf{S}\times\mathbf{S}$ into the reals is suitable as an extended conditional belief function on **S** just in case the corresponding function P' on $\mathscr{E}$ such that

$$P'(B/A) = P(\mathbf{b}/\mathbf{a})$$

satisfies 1–4 with respect to $\mathscr{E}$. On this formulation there is a representation theorem.

THEOREM 2.3: P is suitable as an extended conditional belief function on **S** iff P is a Popper function on S.

Proof. Just as in theorems 1 and 2 except that compactness now follows from compactness of truth functional logic.■

Stalnaker's result is a general representation theorem for Popper functions. Any minimal algebra **F** can have a truth valuation put on it. The truth valuations provide a compact field within which to construct $K(P'_A)$. From Stalnaker's theorem we see that putting 1–4 on these $K(P'_A)$ will insure that P is a Popper function.

This also indicates that when $\mathscr{E}$ is an algebra with structure we care about there will be many Popper functions on $\mathscr{E}$ that ignore and even clash with the structural properties of $\mathscr{E}$. This is an obvious result of the fact that the Popper function is not based on the structure of $\mathscr{E}$ but induces whatever structure it needs onto $\mathscr{E}$. Because we cared about the structure of the proposition space we made belief functions responsive to it by defining $K(P_A)$ on the proposition field itself. This does not mean, however, that Popper's important theorem about Popper function induced structure is not useful for representation of belief. In fact the equivalence classes induced by P play a very important role.

3. *Conceptual Frameworks*

Suppose that P is an extended belief function on a field $\mathscr{E}$ of subsets of T. The algebra of equivalence classes induced by P on $\mathscr{E}$ corresponds to an important conceptual structure for a P-agent. The P-equivalence relation,

$$A \underset{P}{\sim} B \quad \text{iff} \quad P(A/C) = P(B/C) \quad \text{for all} \quad C,$$

holds between just those pairs of propositions that, for a P-agent, are not distinguishable by means of any possible evidence. No assumption whatever counts as more evidence for one than the other.

The maximum element of $\mathscr{E}//p$ is the set of all A such that $P(A/\bar{A})=1$. By remark 1.3i, $P(A/\bar{A})=1$ if and only if $P(A/C)=1$ for all C. Thus, A is P-valid just in case a P-agent would accept A relative to any assumption C. If A is P-valid then a P-agent will count nothing as evidence against A. The P-valid propositions can be regarded as postulates of the agent's conceptual framework.

The basic constraints on extended belief functions insure that P-validity must conform to semantical possibility relative to $\mathscr{E}$. Let

$$\Delta^*(P) = \{A : P(A/\bar{A}) = 1\}$$

and

$$K^*(P) = \bigcap \{A : P(A/\bar{A}) = 1\}$$

From the axioms on Popper functions we have

$$\Delta^*(P) \subseteq \Delta(P_A) \quad \text{for all} \quad A$$

and, thus,

$$K(P_A) \subseteq K^*(P) \quad \text{for all} \quad A.$$

From this it follows that

$$K^*(P) \neq \emptyset$$

by constraint 2ii on belief functions. If $K^*(P)=\emptyset$ then $K(P_A)=\emptyset$ for all A which violates 2ii.

The other natural assumption about $K^*(P)$ is that

$$K^*(P) = T$$

so that $P(A/\bar{A})=1$ only if A is true in every possible world of $\mathscr{E}$. This does not follow from the constraints on extended conditional belief functions, and it should not be added.[17] One of the beauties of the Popper function representation is that part of the conceptual framework can be read off from the belief function. Nothing is lost by letting the structure of $\mathscr{E}$ be less specific than that given by $\mathscr{E}//P$. By letting $K^*(P)$ be less than all of T we can impose whatever specific meaning postulates that may be peculiar to the agent's conceptual framework. Finally, when $K^*(P)$ is allowed to be less than all of T there is room for the kind of conceptual change would correspond to giving up some meaning postulate.

Let h map $\mathscr{E}$ into subsets of $K^*(P)$ so that for all A,

$$h(A) = A \cap K^*(P)$$

and let $\mathscr{E}_P^*$ be the range of h.

Remark 3.1. $\mathscr{E}_P^*$ is isomorphic to $\mathscr{E}//P$.

Proof: What we must show is that

$$\text{(i)} \qquad h(A) = h(B) \quad \text{iff} \quad P(A/C) = P(B/C) \quad \text{for all} \quad C.$$

Suppose $h(A)=h(B)$ and note that $K(P_C) \cap A = K(P_C) \cap B$, since $K(P_C) \subseteq K^*(P)$. Thus, every bet for A relative to $K(P_C)$ has exactly the same consequences as a bet for B relative to $K(P_C)$. Suppose $P(A/C)=$ $=P(B/C)$ for all C

$$A \cap \bigcup_{C \in \mathscr{E}} K(P_C) = B \cap \bigcup_{C \in \mathscr{E}} K(P_C)$$

Therefore $A \cap K^*(P) = B \cap K^*(P)$.

The reduced field $\mathscr{E}_P^*$ can do the job of $\mathscr{E}$ in that every $\mathscr{E}_{P_A}$ is isomorphic to a quotient algebra of $\mathscr{E}//P$ modulo the filter corresponding to $\Delta(P_A)$. If $P(A/\bar{A}) \neq 1$ then $\mathscr{E}_{P_A}$ cannot do this job, because for any B such that $P(B/A)=0$, $K(P_A) \cap K(P_B) = \emptyset$.

4. *Expected Utility and Measurement*

The basic assumption here is that if the agent can conceive of evidence that would support A then he ought to be able to make hypothetical choices relative to the assumption that A holds. Thus, I assume that whenever $P(\bar{A}/A) \neq 1$ the agent has expected utilities defined for $K(P_A)$.

One of the standard puzzles about counterfactual assumptions is that there are often incompatible alternative ways to alter one's background knowledge to accommodate the assumption. In the framework of partial belief this problem is not crucial. As we have been expounding the heuristics for it, $\Delta(P_A)$ will include only those propositions that the agent is sure he ought to hold if he were to accept A. The agent will not choose between alternatives unless he is sure of one of them. This in no way affects the construction $K(P_A)$. There is no need to choose one of them since the agent can utilize finer partitions of $K(P_A)$ to define his expected utilities for it.

No general method for constructing exactly what $K(P_A)$ ought to be, given what $K(P_T)$ and A are, has been provided. What has been done is to provide constraints 1–4 on what $K(P_A)$ constructions are permissible. These constraints have been shown to be sufficient to have P a Popper function.

Though I do not present the details here, the outline of a treatment of expected utility appropriate to the Popper function representation of belief is not difficult. Expected utility is relativized to assumptions. Where $B_0 \ldots B_{n-1}$ is a partition of $K(P_A)$ and $a_0 \ldots a_{k-1}$ are hypothetical acts to be decided upon and for all a_j and B_i $P(B_i/A)$ is independent of a_j; then

$$E_A(a_j) = \sum_{i<n} U(a_j B_i) \cdot P(B_i/A).$$

More generally when $B_0 \ldots B_{n-1}$ is a partition of $K^*(P)$

$$E_A(a_j) = \sum_{i<n} U(a_j B_i \cap K(P_A)) \cdot P(B_i/A).$$

Appropriate modifications for any of the present axiomatic treatments of expected utility should be fairly routine. The main change is that $K(P_A)$ rather than A is the appropriate proposition to relativize to when accessing $E_A(a_j)$.

Measurement of counterfactual conditional probabilities is made by relativizing the Ramsey measuring choice to $K(P_A)$.

Assume that it were that A, then relative to this assumption choose between

(a) receive prize x if B and nothing if not
(b) receive prize x if any one of m of the n random outcomes of gambling device comes up, nothing otherwise.

This choice involves making a hypothetical assumption that may conflict with what the agent accepts, but the ordinary Ramsey measurement may do so as well. For many propositions I am quite sure that there is no way of making sure that receipt of the prize attaches to the truth of B. I doubt whether the counterfactual choices my measurements would require need be any worse off than some of those Savage's framework would require.

Once again, the most important idea supporting the extension of expected utility reasoning to cover assumptions that conflict with what the agent accepts is this: To the extent that the agent can conceive of possible evidence that would lead him to accept something that conflicts with what he now accepts he ought to be able to make hypothetical judgments about what actions would be appropriate if it were the case. In a nut shell, to the extent the agent seriously allows for the possibility that something he accepts may be mistaken he should be able to plan for such contingencies.

III. COUNTERFACTUALS AND ITERATED CONDITIONALIZATION

1. *Iterative Probability Models*

One motivation for extending belief functions from classical probabilities to Popper functions is to be able to revise previously accepted evidence. When belief functions are represented only as classical probabilities previously accepted evidence cannot be revised by conditionalization on later inputs.

If P is only a classical probability and $P(A)=1$ then $P(A/C)=1$ for every C such that $P(C)>0$ and $P(A/C)$ is not defined when $P(C)=0$. If P is a Popper function $P(A/C)$ can be non-trivially defined even when $P(C)=P(C/T)=0$. This goes part of the way toward a solution.

It is not hard to see, however, that extension to Popper functions cannot be the whole solution. Even though P_C may be well defined when $P(C)=0$ so that the agent can shift his absolute belief function from P_T to P_C many of the new conditional belief assignments are not specified. For any B such that $P_C(B)=0$, no values for $P_C(H/B)$ are specified. Failure to specify these conditional beliefs can result in later failure to specify absolute degrees of belief. If after having found reason to accept C one were to later find reason to accept some further proposition B such that $P_C(B)=0$, then his appropriate new absolute belief function would not be specified. Introducing the Popper function representation does allow corrigibility; but, without some further apparatus, this only extends as far as one correction.

The problem with iterated shifts would be solved if one were able to specify not only the new absolute belief function P_C but also the rest of the values for an appropriate new Popper function $P_{\langle C\rangle}$. If such a $P_{\langle C\rangle}$ is defined for each extended belief function P and proposition C such that $P(\bar{C}/C)\neq 1$, then the appropriate shifts in belief as one accepts new evidence can be iterated even when there are iterated clashes of evidence. The shift from P to $P_{\langle C\rangle}$ will give all the P_C values so that $P_C(B)=$ $=P_{\langle C\rangle}(B/T)$, and it will define $P_{\langle C\rangle}(H/B)$ when $P_{\langle C\rangle}(B/T)=0$. Upon being confronted with B as a new input, a new shift from $P_{\langle C\rangle}$ to $P_{\langle C\rangle\langle B\rangle}$ is made. In order to achieve a representation with this kind of richness additional apparatus is needed.

In order to see how to go about doing this, it will be helpful to reconsider Carnap's confirmation and credence functions. If K expresses everything a rational agent accepts, then, according to Carnap, his credence function $\mathscr{C}_{(K)}$ should conform to the confirmation function $\mathscr{C}$ so that

$$\mathscr{C}_{(K)}(B/A)=\mathscr{C}(B/K\cap A).$$

Where $K\cap A$ is empty, $\mathscr{C}_{(K)}$ is undefined.

Let us investigate what happens when we represent rational credence by means of Popper functions. Since we shall want to speak of $\Delta(P_{\langle C\rangle}, A)$ we shift our notation from $\Delta(P_A)$ to $\Delta(P, A)$. We understand $\Delta(P, A)=$ $=\{B: P(B/A)=1\}$ to be the set of propositions a P-agent would accept were he to minimally revise his beliefs to accept A. If the set $K(P, A)=$ $=\cap\Delta(P, A)$ is a member of $\mathscr{E}$, then it is a single proposition expressing the total evidence accepted after the shift to accept A. This suggests that the appropriate relationship between the Popper function representation of rational credence and its corresponding confirmation function ought to be

$$P(B/A)=\mathscr{C}(B/K(P, A))$$

provided $K(P, A)$ is non-empty. When $K(P, A)$ is empty, $P(\bar{A}/A)=1$ and A is regarded as absurd A is not a possible candidate for being accepted as a new input. Thus, there is no loss from having $\mathscr{C}$ undefined for the empty proposition $\emptyset$.

Our constraints on $\mathscr{E}$ do not ensure that $K(P, A)$ will be a member of $\mathscr{E}$. If $K(P, A)$ were denumerable and $\mathscr{E}$ were not closed under denumerable intersections, then we could have it that $K(P, A)$ is not in $\mathscr{E}$. We shall assume that $\mathscr{E}$ is closed under the formation of $K(P, A)$ for every A, and that $\mathscr{C}$ is defined on $\mathscr{E}$.

The connection between P and $\mathscr{E}$ allows $P_{\langle C\rangle}(B/A)$ to be specified in terms of $K(P_{\langle C\rangle}, A)$ so that

$$P_{\langle C\rangle}(B/A) = \mathscr{C}(B/K(P_{\langle C\rangle}, A)).$$

We are assuming that $\mathscr{C}(B/K(P_{\langle C\rangle}, A))$ represents the degree of belief in B appropriate to a rational agent with $K(P_{\langle C\rangle}, A)$ as his total evidence, and that $K(P_{\langle C\rangle}, A)$ is the total evidence accepted by a $P_{\langle C\rangle}$ agent who minimally revises his beliefs to accept A.

This goes part of the way toward solving the shifting problem because $K(P_{\langle C\rangle}, A)$ can be defined in terms of acceptance alone. No values of $P_{\langle C\rangle}$ except those where $P_{\langle C\rangle}(B/A)=1$ need be considered. If $\mathscr{E}$ were closed under a binary propositional function f such that

$$P_{\langle C\rangle}(f(AB)) = 1 \quad \text{iff} \quad P_{\langle C\rangle}(B/A) = 1$$

then $K(P_{\langle C\rangle}, A)$ could be defined in terms of P. We have $P_{\langle C\rangle}(A)= =P(A/C)$ so that

$$\Delta(P_{\langle C\rangle}, A) = \{B: P(f(AB)/C) = 1\},$$

and

$$K(P_{\langle C\rangle}, A) = \bigcap \{B: P(f(AB)/C) = 1\}.$$

Given all this, $P_{\langle C\rangle}$ could be defined in terms of $\mathscr{C}$ and P.

One paradigm for counterfactual conditionals is characterized by acceptability conditions that meet the requirements we want for f. I call this the Ramsey test paradigm because it is characterized by a version of Ramsey's test for acceptability of hypotheticals (Ramsey [47] p. 24. Robert Stalnaker is responsible for the specific version of the test. He uses it to characterize the use of conditionals he intends his theory to explicate (Stalnaker [53]. The test is summed up in the following slogan. Accept $A \Box\!\!\rightarrow B$ (the conditional with antecedent A and consequent B) if and only if the minimal revision of your system of beliefs needed to accept A also requires accepting B. On the Popper function representation of belief, this comes to:

$$\text{Accept } A \Box\!\!\rightarrow B \quad \text{iff} \quad K(P, A) \subseteq B.$$

Any conditional with these acceptability conditions for rational belief functions will satisfy the constraints on f.

Let us augment $\mathscr{E}$ to include closure under $A \Box\!\!\rightarrow B$. We assume that both P and $\mathscr{C}$ are defined on the augmented $\mathscr{E}$. For each C such that $P(\bar{C}/C) \neq 1$ we have, for all A and B

I. (i) $\Delta(P_{\langle C\rangle}, A) = \{B: P(A \Box\!\!\rightarrow B/C) = 1\}$

(ii) $K(P_{\langle C\rangle}, A) = \bigcap \Delta(P_{\langle C\rangle}, A)$

(iii) $P_{\langle C\rangle}(B/A) = \mathscr{C}(B/K(P_{\langle C\rangle}, A))$, if $K(P_{\langle C\rangle}, A) \neq \emptyset$ and $P_{\langle C\rangle}(B/A) = 1$ if $K(P_{\langle C\rangle}, A) = \emptyset$.

We are interested in iterated shifts. Write '$P_{\langle C\rangle_n}$' for '$P_{\langle C_0\rangle \ldots \langle C_{n-1}\rangle}$' and let '$P_{\langle C\rangle_0}$' denote P. A Popper function adequate for iterated shifting must have $P_{\langle C\rangle_{n+1}}$ adequate for shifting whenever $P_{\langle C_n\rangle}(\bar{C}_n/C_n) \neq 1$. To this end we give the following definition of an iterative probability model relative to $\mathscr{C}$.

II. P is $I_p(\mathscr{C}\mathscr{E})$ (P is an iterative probability model for $\mathscr{E}$ relative to $\mathscr{C}$).

iff P is a Popper function on $\mathscr{E}$, $\mathscr{C}$ is a confirmation function that agrees with P, and for any $n+1$ length sequence C of propositions such that $P(\bar{C}_n/C_n) \neq 1$

(a) $P_{\langle C\rangle_n}(A \Box\!\!\rightarrow B/T) = 1$ iff $P_{\langle C\rangle_n}(B/A) = 1$

(b) $P_{\langle C\rangle_n}$ is a Popper function on $\mathscr{E}$

(c) $P_{\langle C\rangle_{n+1}}(B/T) = P_{\langle C\rangle_n}(B/C_n)$

One least elementary remark.

(d) If P is $I_p(\mathscr{E})$ and $P_{\langle C\rangle_n}(\bar{C}_n/C_n) \neq 1$ then $P_{\langle C\rangle_{n+1}}$ is $I_p(\mathscr{E})$.

Each non-trivial $P_{\langle C\rangle_n}$ is itself an iterative probability model.

In the presence of (a) the requirement that $P_{\langle C\rangle_n}$ be a Popper function is equivalent to three constraints on $P_{\langle C\rangle_n}$ assignments to conditionals.

R.1.1. $P_{\langle C\rangle_n}$ is a Popper function iff

(1) $P_{\langle C\rangle_n}(A \Box\!\!\rightarrow A) = 1$

(2) $P_{\langle C\rangle_n}(A \Box\!\!\rightarrow B) \neq 1$ for some A and B, and

(3) If $P_{\langle C\rangle_n}(A \Box\!\!\rightarrow \bar{B}) \neq 1$ then $P_{\langle C\rangle}(AB \Box\!\!\rightarrow D) = 1$ iff $P_{\langle C\rangle_n}(A \Box\!\!\rightarrow (B \supset D)) = 1$.

Proof: The most important step is that (3) is equivalent to

(4) $K(P_{\langle C\rangle_n}, AB) = K(P_{\langle C\rangle_n}, A) \cap B,$
provided $K(P_{\langle C\rangle_n}, A) \cap B \neq \emptyset.$

That (3) and (4) are, thus, equivalent follows quite straight-forwardly from definition I and (a). To show that (1)–(3) are sufficient to have $P_{\langle C\rangle_n}$ be a Popper function we need only note that (1) and (2) together with definition I and (a) yield that $P_{\langle C\rangle_n}$ satisfies constraints c1–c3 on extended conditional belief functions. Since (4) is just constraint c4 applied to $P_{\langle C\rangle_n}$ theorem II 2.1 yields the desired result. To show that if $P_{\langle C\rangle_n}$ is a Popper function then (1)–(3) are satisfied note that by (a), (1) and (2) follow trivially from the basic constraints on Popper functions. Since c4, also, holds for Popper functions we have (3) as well.■

2. *Ip-validity and Conditional Logic*

Let us investigate validity relative to iterative probability models. Assume that $\mathscr{E}$ is a countable field of propositions closed under conditionals, so that $A \Box\!\!\rightarrow B$ is in $\mathscr{E}$ for every A and B in $\mathscr{E}$. We define iterative probability validity for A in $\mathscr{E}$.

IV. A is Ip-valid iff for every iterative probability model P on $\mathscr{E}$ $P(A/C)=1$ for every C in $\mathscr{E}$.

For a given probability model P we already had a notion of P-validity in that A is P-valid just in case $P(A/C)=1$ for all C. Our stronger notion of Ip-validity is the obvious one that A be P-valid relative to every iterative probability model P.

The following axiomatization characterizes validity for David Lewis' basic logic **VC** for counter-factual conditionals.[18] Where A, B, C and D are any propositions in $\mathscr{E}$ the following are the rules and axioms:

Rules (1) Modus Ponens

(2) Deduction within conditionals: for any $n \geqslant 1$,

$$\frac{\vdash (\bigcap_{i<n} A_i) \supset B}{\vdash (\bigcap_{i<n} (C \Box\!\!\rightarrow A_i)) \supset (C \Box\!\!\rightarrow B)}$$

Axioms (1) All truth functional tautologies

(2) $A \Box\!\!\rightarrow A$

(3) $(\bar{A}\Box\!\!\rightarrow A)\supset(B\Box\!\!\rightarrow A)$
(4) $\overline{(A\Box\!\!\rightarrow\bar{B})}\supset(((AB\Box\!\!\rightarrow D)\equiv(A\Box\!\!\rightarrow(B\supset D)))$
(5) $(A\Box\!\!\rightarrow B)\supset(A\supset B)$
(6) $AB\supset(A\Box\!\!\rightarrow B)$

The class $V(T)$ of **VC**-valid propositions of $\mathscr{E}$ is the smallest subset of $\mathscr{E}$ containing every instance of each axiom and closed under the rules. The axiomatization also characterizes valid consequence for **VC**-logic. Suppose Δ is a subset of $\mathscr{E}$. The class $V(\Delta)$ of **VC**-valid consequences of Δ is the smallest subset of $\mathscr{E}$ which includes $\Delta\cup V(T)$ and is closed under Modus Ponens. **VC**-consistency is defined in the usual way, i.e. Δ is **VC**-consistent just in case there is no B such that B and $\bar{B}$ are both in $V(\Delta)$.

Theorem 2.1 A is Ip-valid iff A is **VC**-valid.

Proof. From right to left. We show that the **VC**-axioms are Ip-valid and that the **VC**-rules preserve Ip-validity. Thus, we show the soundness of the **VC**-axiomatization for Ip-models. This is facilitated by Remark 1.3 on Popper functions.

1.3 (i) $P(A/\bar{A})=1$ iff $P(A/C)=1$ for all C
(ii) $P(A/\bar{A})\neq 1$ iff $P(A/C)=1$ for some C.

If we cannot consistently assume that $P(A/C)=0$ for some Ip-model P and proposition C then A is Ip-valid. By remark d (this section) we have that $P_{\langle C\rangle}$ is an Ip-model if P is and $P(\bar{C}/C)\neq 1$. Therefore, to show A is Ip-valid it suffices to show that there is no Ip-model P such that $P(A)= =P(A/T)=0$.

That modus-ponens preserves Ip-validity follows trivially from the fact that when $P(\bar{A}\cup B)=1$ then $P(A)\leqslant P(B)$. Consider Rule 2: Assume $(\bigcap_{i<n}A_i)\supset B$ is Ip-valid and that P is an Ip-model such that

$$P((\bigcap_{i<n}(C\Box\!\!\rightarrow A_i))\supset(C\Box\!\!\rightarrow B))=0.$$

We have $P(A_i/C)=1$ for all A_i and, thus, that $P(\bigcap_{i<n}A_i/C)=1$. We also have $P(B/C)=0$. But, by the Ip-validity of $(\bigcap_{i<n}A_i)\supset B$ we also have $P(B/C)=1$ which is impossible. The Ip-validity of the axioms is established similarly. This completes showing the soundness of the **VC**-axiomatization for Ip-validity.

It is worth pointing out, however, that some of Lewis' axioms correspond directly to salient features of Ip-models. Axiom 3, $(\bar{A}\Box\!\!\rightarrow A)\supset\supset(B\Box\!\!\rightarrow A)$, corresponds to R 1.3 (i) (above) and axiom 1, $A\Box\!\!\rightarrow A$, corresponds to the Popper function constraint that $P(A/A)=1$. The most striking case is axiom 4,

$$\overline{(A\Box\!\!\rightarrow \bar{B})}\supset(((AB\Box\!\!\rightarrow D)\equiv(A\rightarrow(B\supset D))),$$

which corresponds to constraint (3) on acceptability of conditionals,

$$\begin{aligned}&\text{If}\quad P(A\Box\!\!\rightarrow\bar{B})\neq 1,\quad\text{then}\\&P(AB\Box\!\!\rightarrow D)=1\quad\text{iff}\quad P(A\Box\!\!\rightarrow(B\supset C))=1,\end{aligned}$$

and, thus, also to the main $K(P_A)$ condition,

$$K(P, AB)=K(P, A)\cap B,\quad\text{provided}\quad K(P, A)\cap B\neq\emptyset,$$

on extended conditional belief functions. Lewis apologizes for having to use such a long and unintuitive axiom. There is some interest in seeing that in the IP-framework this axiom corresponds directly to a very natural constraint on acceptance.

Let us turn now to the other half of the theorem and show that each Ip-valid A is also **VC**-valid. What we show here is that the **VC**-axiomatization is complete with respect to Ip-validity. If A is Ip-valid then $P(\bar{A})=0$ for every Ip-model P. Therefore, if for each **VC**-consistent proposition A there is an Ip-model P such that $P(A)=1$ then every Ip-valid proposition is a theorem of the **VC**-axiomatization. Thus, the standard Henkin-Lindenbaum procedure for showing completeness is applicable.

Suppose that A is **VC**-consistent (i.e. $\{A\}$ is **VC**-consistent). Lindenbaum's lemma holds for **VC**-consistency (Lewis [37] p. 125). Therefore, there is a maximal **VC**-consistent subset Δ of $\mathscr{E}$ such that $A\in\Delta$. We use Δ to define that part of P where $P(B/A)=1$ so that for every A, B in $\mathscr{E}$

$$P(B/A)=1\quad\text{iff}\quad A\Box\!\!\rightarrow B\in\Delta.$$

One **VC**-theorem is $(T\Box\!\!\rightarrow A)\equiv A$. Since $A\in\Delta$ so is $T\Box\!\!\rightarrow A$. Thus, $P(A)=P(A/T)=1$. For any sequence $C\in\mathscr{E}^n$ and $A\in\mathscr{E}$ we have

$$K(P_{\langle C\rangle n}, A)=\cap\{B:(C_0\Box\!\!\rightarrow\ldots(C_{n-1}\Box\!\!\rightarrow B))\ldots)\in\Delta\}$$

and

$$K(P_{\langle C\rangle_n}, A)=\emptyset\quad\text{iff}\quad P_{\langle C\rangle_n}(B/A)=1\quad\text{for all}\quad B.$$

These follow straight forwardly from the basic properties of **VC**-consistency and definition I.

Let $\mathscr{E}^*$ be the σ-field generated by $\mathscr{E}$. Since $\mathscr{E}^*$ is countable there exist classical measures M on $\mathscr{E}^*$ that assign $M(A)>0$ to every A in $\mathscr{E}^*$ such that $A \neq \emptyset$. Any such M generates a regular confirmation function $\mathscr{C}$ on $\mathscr{E}^* \times \mathscr{E}^* - \{\emptyset\}$ where

$$\mathscr{C}(H/E) = M(H \cap E)/M(E).$$

Let $\mathscr{C}$ be any such confirmation function on $\mathscr{E}^*$.

The rest of the P-values can be defined relative to $\mathscr{C}$ in the manner of definition I. For all A, B in $\mathscr{E}$ let

$$P(B/A) = \mathscr{C}(B/K(P, A))$$

provided $K(P, A) \neq \emptyset$. What remains is to show that P is an Ip-model for $\mathscr{E}$ relative to $\mathscr{C}$. The basic constraint that $P(A \Box\!\!\rightarrow B/T)=1$ iff $P(B/A)=1$ results from the **VC**-validity of

(T1) $\quad (T \Box\!\!\rightarrow B)) \equiv (A \Box\!\!\rightarrow B).$

One of the **VC**-axioms (VC2) is $A \Box\!\!\rightarrow A$. Therefore $A \Box\!\!\rightarrow A \in \Delta$, for every A in $\mathscr{E}$ and

(1) $\quad P(A/A) = 1 \quad \text{for all} \quad A.$

The maximal consistent set Δ cannot have $T \Box\!\!\rightarrow \bar{T}$ in it or it would not be **VC**-consistent. Therefore,

(2) $\quad P(B/A) \neq 1 \quad \text{for some} \quad A \text{ and } B.$

Finally Lewis' axiom

(VC4) $\quad \overline{(A \Box\!\!\rightarrow \bar{B})} \supset (((AB \Box\!\!\rightarrow D) \equiv (A \Box\!\!\rightarrow (B \supset D)))$

insures that

(3) $\quad$ If $P(A \Box\!\!\rightarrow \bar{B}) \neq 1$ then
$P(AB \Box\!\!\rightarrow C) = 1 \quad \text{iff} \quad P(A \Box\!\!\rightarrow (B \supset D)) = 1.$

Therefore since the basic constraint holds for P remark R.1.1 yields that P is a Popper function.

Assume C is a sequence of propositions of length $n+1$, and that $P_{\langle C \rangle_n}(\bar{C}_n/C_n) \neq 1$. The following derived inference rules hold for any sequence C.

(DR1) If $A \in V(T)$ then
$(C_0 \Box\!\!\rightarrow \ldots (C_n \Box\!\!\rightarrow A) \ldots) \in V(T)$.

(DR2) If $(A \equiv B) \in V(T)$ then
$(C_0 \Box\!\!\rightarrow \ldots (C_{n-1} \Box\!\!\rightarrow A) \ldots) \equiv (C_0 \Box\!\!\rightarrow \ldots (C_{n-1} \Box\!\!\rightarrow B) \ldots$
$\ldots) \in V(T)$.

When (DR2) is applied to T1) we have

(a) $P_{\langle C \rangle_{n+1}}(A \Box\!\!\rightarrow B/T) = 1$ iff $P_{\langle C \rangle_{n+1}}(B/A) = 1$.

When (DR2) is applied to

(T2) $(C_n \Box\!\!\rightarrow (T \Box\!\!\rightarrow B)) \equiv (T \Box\!\!\rightarrow (C_n \Box\!\!\rightarrow B))$

we have

$$K(P_{\langle C \rangle_{n+1}}, T) = K(P_{\langle C \rangle_n}, C_n)$$

which yields clause

(c) $P_{\langle C \rangle_{n+1}}(B/T) = P_{\langle C \rangle_n}(B/C_n)$.

By (c) and the assumption we have $P_{\langle C \rangle_{n+1}}(\bar{C}_n/T) \neq 1$ and thus that (2) holds for $P_{\langle C \rangle_{n+1}}$. That (1) holds for $P_{\langle C \rangle_{n+1}}$ follows from applying (DR1) to axiom VC-2. Similarly, that (3) holds follows from applying (DR1) to axiom (VC-4). Thus, we have shown that P is an Ip-model for $\mathscr{E}$ relative to $\mathscr{C}$.■

3. *Construction of Non-Trivial Ip-Models*

In the course of the completeness proof we showed that the constraints on Ip-models are consistent by constructing one, but, we did not show that there exist significantly non-trivial Ip-models. The Ip-model we constructed need not have more than the two values 0 and 1. In fact, the construction used can be generalized to produce Ip-models of significant complexity.

Suppose that $\mathscr{E}$ is a countable Lewis σ-field. We say that $\mathscr{E}$ is a Lewis field (σ-field) just in case

(1) $\mathscr{E}$ is a field (σ-field) of subsets of a non-empty set T.

(2) $\mathscr{E}$ is closed under a binary operator $A \Box\!\!\rightarrow B$ which satisfies the **VC**-axiomatization, (i.e. every instance of a **VC**-valid scheme equals T).

Since $\mathscr{E}$ is a countable σ-field it is an atomic boolean algebra where each atom a is the intersection $\bigcap u$ of an ultrafilter u of $\mathscr{E}$.

The class of regular confirmation functions on $\mathscr{E}$ is the class of conditional probability functions given by classical probabilities on $\mathscr{E}$ which assign positive probability to every atom. When the number of atoms is countable one can use classical probability measures to construct σ-additive regular confirmation functions. An elementary result in measure theory is the following construction for each such probability measure. Let g be a function which assigns positive real numbers to atoms of $\mathscr{E}$ so that the sum for all atoms is 1. The function m on $\mathscr{E}$ defined so that

(3) $$m(A) = \sum_{a \subseteq A} g(a)$$

for atoms a is a classical probability measure on $\mathscr{E}$ which assigns positive probability to each atom. Moreover, every such measure m is generated by some such g. Thus, a very large variety of regular confirmation functions can be constructed.

For any **VC**-consistent set s of propositions in $\mathscr{E}$ the set **VC**(s) of consequences of s is a filter in $\mathscr{E}$ and every filter in $\mathscr{E}$ is **VC**(s) for some **VC**-consistent set of propositions. For each filter f and regular confirmation function $\mathscr{C}$ there is an Ip-model on $\mathscr{E}$ relative to $\mathscr{C}$.

THEOREM 3.1: If $\mathscr{C}$ is a regular confirmation function covering $\mathscr{E}$ and f is a filter of $\mathscr{E}$ then the function P defined on $\mathscr{E} \times \mathscr{E}$ so that

$$P(B/A) = \mathscr{C}(B/\cap \{D:(A \,\Box\!\!\rightarrow D) \in f\}), \quad \text{provided}$$
$$\cap \{D:(A \,\Box\!\!\rightarrow D) \in f\} \neq \emptyset$$
$$P(B/A) = 1, \quad \text{otherwise}$$

is an Ip-model for $\mathscr{E}$ relative to $\mathscr{C}$.

Proof: The proof is a duplicate of that used in showing that the P constructed in the completeness proof is an Ip-model.■

This theorem shows that non-trivial Ip-models exist. Whatever complexity is built into $\mathscr{C}_{K(P',A)}$ will also characterize P'_A where P' is any $P_{\langle C \rangle_n}$ given by P. There is no difficulty in having P'_A quite rich even though $P'(A) = 0$.

4. *Getting Rid of the Confirmation Function*

In the definition of Ip-model confirmation functions were appealed to. The idea was that there should be some probability function $\mathscr{C}$ such that the agents beliefs were as though they conformed to $\mathscr{C}$ by

(1) $$P(B/A) = \mathscr{C}(B/K(P, A)).$$

Suppose that $\mathscr{E}$ is closed under formation of $K(P_{\langle C\rangle_n}, A)$, and consider the following condition.

(2) $$K(P_{\langle C\rangle_n}, K(P_{\langle C\rangle_{n+1}}, A)) = K(P_{\langle C\rangle_{n+1}}, A).$$

If this holds for every $P_{\langle C\rangle_n}$ then each $P_{\langle C\rangle_{n+1}}$ is definable in terms of $P_{\langle C\rangle_n}$ by

(3) $$P_{\langle C\rangle_{n+1}}(B/A) = P_{\langle C\rangle_n}(B/K(P_{\langle C\rangle_{n+1}}, A)).$$

There is no need to use any confirmation function, since the P values with $K(P, A)$ second arguments do the same job. In Part I it was suggested that credence and confirmation correspond to two different heuristics for conditional probability. Where $P(B/A)$ is the degree of belief in B appropriate to minimally revising to *add* A and nothing further to what one already has, $\mathscr{C}(B/A)$ is the degree of belief appropriate to a minimal revision to *reduce* everything one accepts to just A.

The idea behind (2) is that for $K(P_{\langle C\rangle_n}, A)$ these two procedures should come to the same thing. If $K(P, A)$ is the minimal revision to add A then the minimal revision to add $K(P, A)$ should just be $K(P, A)$ itself. Thus, $\mathscr{C}(B/K(P, A) = P(B/K(P, A))$.

5. *Probability of Conditionals*

Except for the basic constraints on acceptance of conditionals Ip-models leave open what $P(A\Box\!\!\rightarrow B)$ should be. The inspiration for the present treatment was a system of Stalnaker's (Stalnaker [52] pp. 74–79) based on the hypothesis

(SH) $$P(A\Box\!\!\rightarrow B) = P(B/A) \quad \text{all} \quad A \text{ and } B.$$

David Lewis has shown that Stalnaker's system trivialized in that it can take at most only four values. (Lewis [39] pp. 4–7). Lewis' result applies to any system with (SH) together with

(Ip) $$P(A \rightarrow B/C) = P(B/AC) \quad \text{all} \quad A, B \text{ and } C$$

provided $P(AC) > 0$. Since (Ip) holds for Ip-models Lewis' result insures that all significantly non-trivial Ip-models do not satisfy the Stalnaker hypothesis.[19]

6. *Conceptual Change*

With the addition of the conditional operator the agent can use certain propositions to represent that other propositions are postulates of his conceptual framework. By the basic constraint on probability assignments to conditionals

$$(1) \qquad P(\bar{A}\Box\!\!\rightarrow A) = 1 \quad \text{iff} \quad P(A/\bar{A}) = 1$$

Let '$\Box A$' abbreviate $\bar{A}\Box\!\!\rightarrow A$'. The following rule and axioms for this defined necessity operator are Ip-valid.

$$(\mathrm{K1}) \qquad \frac{\vdash A}{\vdash \Box A}$$

$$(\mathrm{m1}) \qquad \Box A \supset A$$

$$(\mathrm{m2}) \qquad \Box(A \supset B) \supset (\Box A \supset \Box B).$$

These together with truth functional tautologies characterize necessity in modal system *M*, the weakest of Kripke's standard modal systems (Kripke [30]). The S4-axiom

$$(\mathrm{S4}) \qquad \Box A \supset \Box\Box A$$

is not Ip-valid, because one can have both

$$P(A/\bar{A}) = 1$$

and

$$P(\bar{A}\Box\!\!\rightarrow A / \overline{\bar{A}\Box\!\!\rightarrow A}) \neq 1.$$

Thus, system *M* is the modal logic that corresponds to *P*-validity in Ip-models.

Having Ip-models where the S4 axiom fails allows for the representation of conceptual change. We represent the minimal revision to add *A* as a new postulate as

$$P_{\langle \Box A \rangle}.$$

We represent the minimal revision to remove postulate *A* from *P*-valid status as

$$P_{\overline{\langle \Box A \rangle}}$$

One interesting property of this is that

$$P_{\overline{\langle \Box A \rangle}}(A) = 1$$

though of course $P_{\overline{\langle \Box A \rangle}}(\Box A) = 0$. This is as it should be.

I think that being able to represent conceptual change in this way is a strong argument in favour of using M-necessity rather than S4 as the necessity induced by the conditional.[20]

7. *Stalnaker vs. Lewis*

Stalnaker's logic of counterfactuals **VC-S** is equivalent to the result of adding a single axiom

$$(7) \qquad \overline{(A \Box\!\!\rightarrow B)} \supset (A \Box\!\!\rightarrow \bar{B})$$

to Lewis' **VC** (Stalnaker [53], Stalnaker and Thomason [54], Lewis [37], [40]). There is now considerable controversy over the merits of Stalnaker's axiom (e.g. Lewis [37], [38], [39], van Fraassen [57], [60], Pollock [42]). If one adds the constraint

$$(7^*) \qquad P(A \Box\!\!\rightarrow B) = 0 \quad \text{iff} \quad P(B/A) = 0$$

to those imposed on Ip-models then (7) would be valid and Stalnaker's logic would capture Ip-validity.

Even without this, Stalnaker's axiom is valid for 2-valued Ip-models. Since $P(B \cup \bar{B}/A) = 1$ we cannot have both $P(B/A) = 0$ and $P(\bar{B}/A) = 0$. Therefore, either $P(B/A) = 1$ or $P(\bar{B}/A) = 1$ and in neither case can (7) be assigned zero

8. *Ellis on the Logic of Subjective Belief*

In a very interesting investigation of the logic of subjective belief Brian Ellis argues that the correct logic of truth for a system of propositions ought to correspond to what would hold in every admissible two-valued probability system for those propositions (Ellis [11] p. 127). He gives the following reasons:

> For if we are certain of the premises of a valid argument, we ought to be certain of its conclusion, and if we are not certain of the conclusion of a valid argument, then we ought not to be certain of all its premises ....
>
> Consequently, if there is any divergence between our logics of truth and certainty, then either something is wrong with our probability theory or with the way we have applied it to the analysis of arguments or something is wrong with our logic of truth.

If Ellis' principle were right Stalnaker's logic would be the correct logic of subjective belief. The argument he gives however, only supports the weaker principle that the truth logic should capture just those arguments where for any admissible probability system where all the premises receive probability 1.0 the conclusion must also receive probability 1.0. This weaker principle is just Ip argument-validity.

Ellis opens his concluding remarks as follows:

> We have as yet no adequate logic of subjective probability. The classical probability calculus is not an adequate logic of subjective probability because
>
> (a) it is not capable of handling subjective probability claims concerning subjunctive conditions, and
>
> (b) it is not strong enough to deal with compound conditionals.

The present system of Ip-models has been constructed to answer just these needs.

9. *Concluding Remarks*

This is a good place to make it clear that Ip models and Ip-validity are not intended as a theory of rational belief change. The Ip-constraints characterize coherence and coherent shifting given an input. What they do not specify is what inputs are rational under what circumstances. Clearly, a full theory of rational belief change would have to include a theory of inputs.

I think that some discussions between Bayesians and classical testing theorists are confused by the fact that the testing theorist is talking about rational inputs while the Bayesian is talking about coherence. The orthodox Baysian may say that no theory of inputs is needed because the only appropriate inputs are observations and they are so obvious as to requires no theory. Taking seriously the idea that observations are fallible, which can be done in the Ip-framework, indicates that some theory of inputs is needed. Working out the details of one might help bring together some of the good points in the two traditions.

Finally, I should like to point out that Ip-models are very much idealizations. No actual agent can be expected to have his belief function defined for all the propositions in a Lewis algebra, nor is any actual agent expected to attain the semantical omniscience built into the characterization of belief functions. There are two comments to be made on this. First the fragment for which an actual agent's belief function is defined can be expected to include conditionals and counterfactual conditions. Second the Ip-model is a constraint on actual rational degrees of belief in much the way that ordinary logic is a constraint on acceptance. To the extent that the belief function is defined and fails to meet the Ip-constraints it is incoherent.

University of Western Ontario

NOTES

[1] Ramsey [46], De Finetti [12] and Savage [49]. This orthodox tradition is a salient subclass of I. J. Good's 46656 (Good [17]) variations on the Bayesian theme. It is characterized by the representation of belief functions as point probability functions and by its emphasis on the role of belief in guiding decisions.

[2] One of the most ingenious alternative approaches is that of Cox and Good where certain modest assumptions about belief require that there exists as probability function representing the beliefs. See Cox [10], Good [15]. For the most explicit treatment of the mathematical details see Aczel [1] pp. 319–24. See Shimony [51] for a recent application of this argument.

[3] Often the field of propositions can be restricted to what Savage calls a small world situation, Savage [49] pp. 87–90. Where each possible world can be regarded as no more than one of the alternative specifications of those factors that would be relevant to the decision problem.

[4] Everything I shall say about expected utility in this paper can be relativized to such situations. In fact I do not believe that the current treatments for cases where $P(A_i)$ depends on a_j are entirely adequate.

[5] The result would be more interesting if it applied to ordinary coherence as well as strict coherence. As we shall see strict coherence is a bit odd anyway.

[6] Originally Carnap's goal was to find constraints that would make $\mathscr{C}$ completely determined by the semantical properties of $\mathscr{E}$. Thus, $\mathscr{C}$ would represent the logical probability function generated by the field of propositions $\mathscr{E}$. This goal of a single logical $\mathscr{C}$-function now seems unattainable and has been largely given up. For a discussion of the changes in Carnap's program from a basically Popperian point of view see Lakatos [32]. The best statement of the new more modest goals of the Carnapian program is Jeffrey [25].

[7] Carnap defines regularity so that certain propositions are exempt. This is a mistake on his part for the following two reasons. First, the strict coherence argument he explicitly claims as the justification for regularity allows no such exceptions. Secondly, violations of full regularity clash with the basic heuristic that guides the program (see text).

[8] Thus, Carnap's objections to acceptance of hypotheses (Carnap [8] pp. 28–31) do not rest on strict coherence. Indeed the discussion of credence suggests strongly that Carnap allows acceptance of evidence claims.

[9] See May and Harper [41] for discussing the minimum change idea together with some metrics and optimization techniques.

[10] If we replace conditional bets by conditional expected utility and the assumption that the bets-beliefs postulate holds for both P_0 and P_1 by the assumption that utility

assignments relative to propositions in $K(P_0) \cap K(P_1)$ remain unchanged, my version of the dutch book argument is transformed into a version of Savage's conditional expected utility argument for conditionalization.

[11] (Popper [43] pp. 318–358). Popper's published work on this subject is given in a series of papers starting in 1938. Most of these papers together with bibliographic material are included in the pages cited.

[12] A minimal algebra is simply a set closed under a unary and a binary relation.

[13] Popper's main axiomatization uses the weaker

a2′. If $P(A/D)=P(B/D)$ for all D, then $P(C/A)=P(C/B)$.

The present a2 is given as an alternative stronger version on p. 335 of Popper [1]. Popper uses the much weaker assumption

$$P(A/A)=P(B/B)$$

in place of al. I use al only to make things more perspicuous.

[14] In the standard Kolmogorov treatment classical probabilities are defined in σ-fields and are σ-additive. If A is a denumerable sequence of pairwise disjoint sets then

$$P\left(\bigcup_{i<\omega} A_i\right)=\sum_{i<\omega} P(A_i).$$

We may, also, have σ-additive Popper functions. Nothing I shall say in the present paper will turn on the difference between finite additively and σ-additivity. In future work developing the measure theory for Popper functions σ-additivity will be important.

A very nice construction for Popper measures by combining even non-denumerably many classical probability measures has been developed by Bas C. van Fraassen [59].

[15] Proof of al. 3 (i): We proceed by first showing a lemma. The proof of this is essentially that given in Popper [1] p. 352. In the present version a2 is not appealed to.

LEMMA. If $P(\bar{C}/C)\neq 1$ then $P(AB/C)+P(A\bar{B}/C)=P(A/C)$.

Assume $P(\bar{C}/C)=1$ and note

(1) $P(B/AC)+P(\bar{B}/AC)=P(C/AC)+P(\bar{C}/AC)$.

holds in case $P(\overline{AC}/AC)=1$ by constraint 2ii and in case $P(\overline{AC}/AC)\neq 1$ by a3. Multiply both sides of 1 by $P(A/C)$.

(2) $P(A/C)\cdot P(B/AC)+P(A/C)\cdot P(\bar{B}/AC)= =P(A/C)\cdot P(C/AC)+P(A/C)\cdot P(\bar{C}/AC)$.

Using a4 on 2 we get:

(3) $P(AB/C)+P(A\bar{B}/C)=P(AC/C)+P(A\bar{C}/C)$.

Coherence (or a4, a5 and a4) yields $P(AC/C)=P(A/C)$, and $P(A\bar{C}/C)=0$.

We turn now to the main result.

R1.3(i) If $P(A/\bar{A})=1$ then $P(A/B)=1$ all B.

Assume $P(A/\bar{A})=1$, and note that $P(A/B)=1$ trivially if $P(\bar{B}/B)=1$ (by constraint 2ii). Assume $P(\bar{B}/B)\neq 1$ and use the lemma.

(1) $P(AA/B)+P(A\bar{A}/B)=P(A/B)$

Use a4 on 1:

(2) $P(A/B)\cdot P(A/AB)+P(\bar{A}/B)\cdot P(A/\bar{A}B)=P(A/B)$.

By a1, a5 $P(A/AB)=1$. Since $P(A/\bar{A})=1$ we have $P(\bar{A}/\bar{A})=1$ and by 2ii $P(BA/\bar{A})=1$ which yields $P(A/\bar{A}B)=1$ by a4. Thus,

(3) $\quad P(A/B)+P(\bar{A}/B)=P(A/B)$

which completes the proof.

[16] Assume $K(P_B)\sim A\neq\phi$. Thus, $P(\bar{B}/B)\neq 1$ and $P(A/B)\neq 0$. To see that $K(P_{AB})\subseteq$ $\subseteq K(P_B)\cap A$ assume $P(C/B)=1$. $P(A\bar{C}/B)=0$, therefore $P(A/B)\cdot P(\bar{C}/AB)=0$ and $P(C/AB)=1$. To see that $K(P_B)\cap A\subseteq K(P_{AB})$ assume $P(C/AB)=1$ and note that $P(\bar{A}\cup C/B)=1$.

[17] Hughes Leblanc attempted to show that Popper functions on sentences are equivalent to a natural extension of Carnap's treatment of confirmation functions for sentences (Leblanc [33]). The following definitions of a deducibility relation $\vdash$ and extended confirmation function $\mathscr{C}$ are equivalent to ones Leblanc gives (Leblanc [33]).

c0. $\quad\vdash a$ iff a is a tautology
c1. $\quad 0\leqslant\mathscr{C}(b/a)\leqslant\mathscr{C}(a/a)=1$
c2. $\quad$ If $\vdash a\equiv b$ and $\vdash c\equiv d$ then $\mathscr{C}(a/c)=\mathscr{C}(b/d)$
c3. $\quad$ If not $\vdash\bar{a}$ then $\mathscr{C}(\bar{b}/a)=1=\mathscr{C}(b/a)$
c4. $\quad\mathscr{C}(ab/c)=\mathscr{C}(a/c)\cdot\mathscr{C}(b/ac)$

Leblanc gives the following axioms for Popper functions

a1. $\quad 0\leqslant P(b/a)\leqslant P(a/a)=1$
a2. $\quad$ If $a\sim b$ then $P(c/a)=P(c/b)$
a3. $\quad$ If $P(c/a)\neq 1$ then $P(\bar{b}/a)=1=P(b/a)$
a4. $\quad P(ab/c)=P(a/c)\cdot P(b/ac)$
a5. $\quad P(ab/c)=P(ba/c)$,

to which he adds

a6. $\quad\vdash_P a$ iff $P(a/\bar{a})=1$.

He then shows

(1) $\quad$ Any $\vdash$ and $\mathscr{C}$ satisfying c0–c4 also satisfy a1–a6.

He also claims to show

(2) $\quad$ Any P and $\vdash_P$ that satisfy a1–a6 also satisfy c0–c4.

The second claim is false, because $\vdash_P$ need not capture only tautologies. This was first pointed out by Stalnaker (Stalnaker [52] footnote p. 70). I include these remarks, because Stalnaker did not indicate how Leblanc went wrong, nor what was actually proved. Indeed one can have a Popper function where $P(a/\bar{a})=1$ or $P(\bar{a}/a)=1$ for all a. Leblanc misleads himself by using the axiomatization that is supposed to recursively define $\vdash$ as simple constraints on $\vdash$.

What Leblanc actually succeeds in proving is that any Popper function P satisfies c1–c4 relative to $\vdash_P$ and that $\vdash_P$ captures at least all tautologies. Furthermore, his proof of the converse is actually a proof of the stronger claim that any $\vdash$ which captures at least all tautologies and $\mathscr{C}$ which satisfies c1–c4 relative to $\vdash$ also satisfy a1–a6.

[18] Lewis [37] p. 132.

[19] After Lewis' trivialization [39], Stalnaker has given up (SH). (Comment delivered by Stalnaker at CPA 1972). Bas van Fraassen, however, has been developing ingenious attempts to circumvent Lewis' results (van Fraassen [57], [58]). These attempts all reject (Ip). Since (Ip) corresponds to iterated conditionalization I think that the price van Fraassen pays to keep (SH) is too high.

[20] This point was the motivation for developing the method of handling universal instantiation for *M*-versions of conditional logic given in Harper [19].
[21] See next section.

BIBLIOGRAPHY

[1] Aczel, J., *Lectures on Functional Equations and Their Applications*, Academic Press, New York, 1966.
[2] Adams, E., 'The Logic of Conditionals', *Inquiry* **8** (1965) 166–197.
[3] Adams, E., 'Probability and the Logic of Conditionals', in *Aspects of Inductive Logic* (J. Hintikka and P. Suppes, eds.), North Holland, Amsterdam, 1966, pp. 265–316.
[4] Adams, E., *The Logic of Conditionals: An Application of Probability to Deductive Logic* (Manuscript).
[5] Adams, E., 'Prior Probabilities and Counterfactual Conditionals', forthcoming in U.W.O. Series.
[6] Carnap, R., 'Inductive Logic and Rational Decisions', in Carnap and Jeffrey [8], pp. 7–31.
[7] Carnap, R., 'A Basic System of Inductive Logic, Part I', in Carnap and Jeffrey [8], pp. 35–165.
[8] Carnap, R. and Jeffrey, R. L. (eds.), *Studies in Inductive Logic and Probability*, Univ. of California Press, Los Angeles, 1971.
[9] Chisholm, R., 'The Contrary-to-Fact Conditional', in *Readings in Philosophical Analysis* (H. Feigl, and W. Sellars eds.), New York, 1949, pp. 482–497.
[10] Cox, R. T., *The Algebra of Probable Inference*, Johns Hopkins Press, Baltimore, 1960.
[11] Ellis, B., 'The Logic of Subjective Probability', *British Journal for the Philosophy of Science* **24** (1973) 125–152.
[12] de Finetti, B., 'Foresight: Its Logical Laws, Its Subjective Sources', in Kyburg and Smokler [31], pp. 93–158.
[13] de Finetti, B., *Probability, Induction and Statistics*, John Wiley and Sons, York, 1972.
[14] de Finetti, B., 'Initial Probabilities: A Prerequisite for any Valid Induction' *Synthese* **20** (1969) 2–16.
[15] Good, I. J., *Probability and the Weighing of Evidence*, Charles Griffin and Company Limited, London 1950.
[16] Good, I. J., 'The Bayesian Influence or How to Sweep Subjectivism Under the Carpet', forthcoming in the proceedings of the 1973 V.W.O. conference on Foundations of Statistics and Statistical Theories.
[17] Good, I. J., 'Explicativity, Corroboration, and the Relative Odds of Hypotheses', forthcomming synthese.
[18] Goodman, N., *Fact, Fiction and Forecast* (paper back ed.) Bobbs Merrill Co., New York, 1965.
[19] Harper, W. L., 'A Note on Universal Instantiation in the Stalnaker-Thomason Conditional Logic', Forthcoming in *The Journal of Philosophical Logic*.
[20] Harper, W. L. and Kyburg, H. E., 'The Jones Case', *British Journal for the Philosophy of Science* **19** (1968) 247–258.
[21] Hintikka, J., *Knowledge and Belief*, Cornell Univ. Press, Ithaca, 1962.
[22] Hintikka, J., *Models for Modalities*, Reidel Publishing Co., Dordrecht, 1969.

[23] Jeffrey, R. C., *The Logic of Decision*, McGraw-Hill, New York, 1965.
[24] Jeffrey, R. C., 'Probable Knowledge', in Lakatos [1], pp. 166–180.
[25] Jeffrey, R. C., 'Carnap's Inductive Logic', *Synthese* **25** (1973) 299–306.
[26] Kemeny, J., 'Fair Bets and Inductive Probabilities', *Journal of Symbolic Logic* **20** (1955) 263–273.
[27] Kolmogovov, A. N., *Foundations of the Theory of Probability* 2nd English ed., Chelsea, New York, 1956.
[28] Krantz, D. H., Luce, R. D., Suppes, P. and Tversky, A., *Foundations of Measurement*, Academic Press, New York, 1971.
[29] Kripke, S., 'Semantical Considerations on Modal Logic', in *Modal and Many-Valued Logics, Acta Philosophica Fennica* **16** (1963) 83–94.
[30] Kripke, S., 'Semantical Analysis of Modal Logics, I', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **9** (1963) 67–96.
[31] Kyburg, H. and Smokler, H. (eds), *Studies in Subjective Probability*, Wiley, New York, 1964.
[32] Lakatos, I., 'Changes in the Problem of Inductive Logic', in Lakatos (ed.): *The Problem of Inductive Logic*, North-Holland, Amsterdam, 1968, pp. 315–417.
[33] Leblanc, H., 'On Requirements for Conditional Probability Functions', *Journal of Symbolic Logic* **25** (1960) 238–242.
[34] Levi, I., 'Probability Kinematics', *British Journal for the Philosophy of Science* **18** (1967) 197–207.
[35] Levi, I., 'If Jones Only Knew More', *British Journal for the Philosophy of Science* **20** (1969), 153–159.
[36] Levi, I., 'On Indeterminate Probabilities', in *Journal of Philosophy* **71** (1974) 391–418.
[37] Lewis, D. K., *Counterfactuals*, Blackwell, London, 1973.
[38] Lewis, D. K., 'Counterfactuals and Comparative Possibility', forthcoming in Hockney, Harper, Fried (eds.), *Contemporary Research in Philosophical Logic and Linguistic Semantics*, U.W.O. Workshop series.
[39] Lewis, D. K., 'Probabilities of Conditionals and Conditional Probabilities' (dittograph).
[40] Lewis, D. K., 'Completeness and Decidability of Three Logics of Counterfactual Conditionals', *Theoria* **37** (1971) 74–85.
[41] May, S. and Harper, W. L., 'Toward an Optimization Theory for Applying Minimum Change Principles in Probability Kinematics', forthcoming in *Foundations and Philosophy of Selected Epistemic Applications of Probability Theory* (W. L. Harper and C. A. Hooker, eds.), U.W.O. Series volume from 1973 International Workshop on Foundations of Probability and Statistics and Statistical Theories of Science.
[42] Pollock, J., *Counterfactuals, Dispositions and Causes* (Manuscript).
[43] Popper, K. R., *The Logic of Scientific Discovery*, Harper Torchbook edition, New York, 1959.
[44] Popper, K. R., *Conjectures and Refutations*, Harper Torchbook edition, New York, 1965.
[45] Popper K. R., *Objective Knowledge*, Oxford, 1972.
[46] Ramsey, F. P., 'Truth and Probability', in Ramsey [47], pp. 156–211.
[47] Ramsey, F. P., *The Foundations of Mathematics*, Routledge & Kegan Paul Ltd., London, 193.
[48] Rescher, N., *Hypothetical Reasoning*, North Holland, Amsterdam, 1964.

[49] Savage, L. J., *The Foundations of Statistics*, John Wiley & Sons, Inc., 1954.
[50] Shimony, A., 'Coherence and the Axioms of Confirmation', *Journal of Symbolic Logic* **20** (1955) 1–28.
[51] Shimony, A., 'Scientific Inference', in *The Nature and Function of Scientific Theories* (ed. by R. Colodny), Univ. of Pittsburgh Press, Pittsburgh, 1970.
[52] Stalnaker, R., 'Probability and Conditionals', *Philosophy of Science* **37** (1970) 64–80.
[53] Stalnaker, R., 'A Theory of Conditionals', *Studies in Logical Theory* (*American Philosophical Quarterly*, Supplementary Monograph Series), Oxford, 1968.
[54] Stalnaker, R. and Thomason, R., 'A Semantical Analysis of Conditional Logic', *Theoria* **36** (1970) 23–42.
[55] Teller, P., 'Conditionalization and Observation', *Synthese* **26** (1973) 218–258.
[56] Teller, P., 'Conditionalization, Observation and Change of Preference', to appear in *Foundations and Philosophy of Selected Epistemic Applications of Probability Theory* (W. L. Harper and C. A. Hooker, eds.).
[57] van Fraassen, B. C., 'Probabilities of Conditionals', forthcoming in the Proceedings of the 1973 V.W.D. conference.
[58] van Fraassen, B. C., 'Notes on Probabilities of Conditionals' (dittograph).
[59] van Fraassen B. C., 'Construction of Popper Probability Functions' (dittograph).
[60] van Fraassen B. C., 'Hidden Variables in Conditional Logic', forthcoming in *Theoria*.
[61] van Fraassen B. C., 'Theories and Counterfactuals', forthcoming in a festschrift for W. Sellars (ed. by H.-N. Castaneda).
[62] Vickers, J. M., 'Probability and Non-Standard Logics', in *Philosophical Problems in Logic* (Karel Lambert, ed.) Reidel Publ. Co., Dordrecht, 1970, pp. 102–120.
[63] Von Neuman, J. and Morgenstern, O., *Theory of Games and Economic Behavior*, Princeton Univ. Press, Princeton, 1953.

Note added in proof:

Robert Stalnaker has shown that Ip-models trivialize so that $P_{\langle A\rangle}(\bar{B}/B)=1$ whenever $P(A)>0$ and $P(B/A)=0$. (Letter to Author.) Theorem 3.1 of Section III is false. The construction only works when f is an ultrafilter, as in the completeness proof.

In order to meet this serious difficulty the system is revised so that a conditional proposition $A\Box\!\!\rightarrow B$ is allowed to vary with changes in the relevant acceptance context. In evaluating $P(A\Box\!\!\rightarrow B/C)$ the relevant acceptance context is $K(P, C)$. In evaluating $P((A\Box\!\!\rightarrow(B\Box\!\!\rightarrow(C\Box\!\!\rightarrow D)))/E)$ the relevant acceptance context varies with nesting to the right so that $K(P, E)$, $K(P_{\langle E\rangle}, A)$ and $K(P_{\langle E\rangle\langle A\rangle}, B)$ are the contexts relevant for the respective nested conditionals.

Having nearness relativized to acceptance contexts in this way promises to be of some interest in understanding conditionals as well as rational belief.