JAMES HAWTHORNE

ON THE LOGIC OF NONMONOTONIC CONDITIONALS AND CONDITIONAL PROBABILITIES[1]

ABSTRACT. I will describe the logics of a range of conditionals that behave like conditional probabilities at various levels of probabilistic support. Families of these conditionals will be characterized in terms of the rules that their members obey. I will show that for each conditional, $\rightarrow$, in a given family, there is a probabilistic support level r and a conditional probability function P such that, for all sentences C and B, '$C \rightarrow B$' holds *just in case* $P[B \mid C] \geqslant r$. Thus, each conditional in a given family behaves like conditional probability above some specific support level.

INTRODUCTION

Those who investigate the logics of indicative and subjunctive conditionals often recognize an acute similarity between logics for these conditionals and the logic of conditional probabilities.[2] In this paper I will investigate the logics of a class of conditionals that behave precisely like conditional probability functions. For each type of conditional that I will investigate, a conditional assertion behaves like the part of a conditional probability function above some specific level of support. That is, for each conditional, $\rightarrow$, there is probabilistic support level r and a conditional probability function P such that, for all sentences C and B, '$C \rightarrow B$' holds *just in case* $P[B \mid C] \geqslant r$. For some of the conditionals under study (those corresponding to a fairly high value for r) an assertion 'if C, then B' means, roughly, 'C supports B' or 'if C, then B is highly plausible'. But, I will explicate the logics of a broad range of probability-like conditionals: conditionals for which '$C \rightarrow B$' means 'if C, then almost certainly B' (corresponding to a probability close to 1), conditionals for which '$C \rightarrow B$' means 'if C, then *just possibly* B' (corresponding to a probability just barely above 0), and conditionals that correspond to the range of support levels in between. All of these conditionals are *nonmonotonic* – i.e. when 'if C, then B' holds, the addition of new information, D, to the antecedent of the conditional may undermine it, so that 'if D and C, then B' may not hold. Nonmonotonic conditionals can play an important role in *defeasible reasoning*, and have been studied by both philosophers and researchers in Artificial Intelligence.[3]

Journal of Philosophical Logic **25:** 185–218, 1996.

In Section 1 I will describe a simple model of defeasible reasoning that illustrates the roles played by conditional probability functions and nonmonotonic conditionals in uncertain inference. This discussion is intended to motivate the formal treatment of conditionals in subsequent sections. In Section 2 I will explicate the *Popper Functions*, a generalization of conditional probability functions. Subsequent sections will show that there is an intimate relationship between Popper Functions and nonmonotonic conditionals. Section 3 establishes the relationship between Popper Functions and a family of conditionals that Lehmann and Magidor (1992) call the *Rational Consequence* relations. We will see that the Rational relations behave like the *probability* 1 parts of the Popper Functions. In Section 4 I will characterize the families of conditionals that behave like conditional probabilities at levels below 1.

1. THE ROLE OF NONMONOTONIC CONDITIONALS IN DEFEASIBLE REASONING SYSTEMS

Most of the inferences we commonly make are *defeasible*. The addition of new information can undermine the support for conclusions that were quite plausible relative to previous information. One way to formalize defeasible reasoning is to employ a formal logic for *nonmonotonic* conditionals, a logic in which conditional statements represent the support (or lack of support) for a hypothesis by available information.

A conditional, $\rightarrow$, is said to be *monotonic* just in case whenever '$C \rightarrow B$' holds, '$D\&C \rightarrow B$' must also hold (for every sentence D) regardless of the content of the new information expressed by D. For monotonic conditionals the addition of more information to the antecedent cannot undermine the support that the antecedent affords the consequent. All conditionals that arise in classical deductive logics are monotonic. Among the monotonic conditionals that occur in the object languages of these logics are the material conditional and all of the various strict conditionals represented in intensional logics. The metalinguistic *logical consequence* relation and its variants (e.g. the logical consequence relations for intuitionistic and relevance logics) may also be regarded as monotonic conditionals.

When a conditional, $\rightarrow$, is *nonmonotonic*, the antecedent of a conditional assertion only tentatively supports the consequent, supports it *ceteris paribus*. New information may defeat the support for a consequent – i.e., although '$C \rightarrow B$' holds, '$D\&C \rightarrow B$' may not hold, and even '$D\&C \rightarrow \neg B$' may hold. The logic of nonmonotonic conditionals is very similar to the logic of conditional probability. Indeed, nonmonotonic conditionals and conditional probabilities play very similar roles in defeasible reasoning.

Consider the role of conditional probability in the reasoning of a Bayesian agent. Suppose P is a probability function that represents an (ideally rational) agent's belief function regarding some domain. The conditional probability '$P[B \mid C] = r$' represents the degree of confidence that the agent should have in the truth of B when C represents all of her present information (or, all that is relevant to B). Similarly, '$P[B \mid D\&C] = s$' represents the degree of confidence that the agent should have in B if $C\&D$ becomes the total evidence available to her. Conditional probability functions are *nonmonotonic* in that the probability value for $P[B \mid C]$ may differ radically from the value for $P[B \mid D\&C]$; even when the former probability is nearly 1, the latter may be much lower, perhaps near 0. For additional information E, the value of $P[B \mid E\&D\&C]$ may again differ widely from the value of $P[B \mid D\&C]$. This kind of nonmonotonicity leads to defeasibility in the agent's beliefs. For, if at first the agent only knows that C, later learns that D, and still later comes to know E, her degree of confidence in B will vary radically.

Conditional probability functions for Bayesian agents are usually treated as meta-linguistic, semantic relationships between object-languages sentences. The *logic* an agent utilizes in making uncertain inferences largely reduces to the application of the axioms of probability theory to deduce the conditional probabilities of some sentences from others. The axioms play the role of semantic rules that constrain probability functions, much as the semantic rules for sentential deductive logics constrain possible truth-value assignments. And just as one may use the semantic rules of sentential deductive logic to compute the truth-values of some sentences from truth-values of other sentences, one may use the axioms of probability theory to compute probabilities of some sentences from the probabilities of other sentences. Thus, the axioms of probability theory play two related roles in probabilistic logic. They represent a sound (and complete) collection of rules through which some probabilistic assertions may be derived from others; and they supply constraints that any total probability function must satisfy if it is to yield a consistent assignment of probabilities to sentences. Although this is all quite obvious where probability is concerned, I have taken the trouble to make it explicit because nonmonotonic conditionals are often treated differently, as part of the object-language. But, the nonmonotonic conditionals I will explicate are best understood as meta-linguistic semantic relationships, rather like conditional probabilities. And the "axioms" and "inference rules" for these conditionals are best understood as semantic rules.

In many common reasoning contexts precise values for conditional probabilities are ill defined or unavailable. In such contexts an agent may make qualitative inferential leaps of confidence as new information becomes available. Rather than infer a probability or degree of belief for a hypothesis, the agent will simply infer that a hypothesis is to be "accepted" or "not accepted" on the present evidence, where "acceptance" is defeasible. Nonmonotonic conditionals provide a way to model belief revision of this kind. The agent's "personal support relation" is represented as a nonmonotonic conditional, $\rightharpoondown$. The total evidence C supports the acceptance of B if the conditional assertion '$C \rightharpoondown B$' holds for the agent's support relation; when '$C \rightharpoondown B$' does not hold, i.e. when '$C \not\rightharpoondown B$', the evidence does not support the acceptance of B. In effect, a nonmonotonic conditional provides a trivalent support function for the agents beliefs. If C is the total relevant evidence, the agent will "accept B" (when '$C \rightharpoondown B$' holds), "accept $\neg B$" (when '$C \rightharpoondown \neg B$' holds), or remain agnostic toward B (when $C \not\rightharpoondown B$ and $C \not\rightharpoondown \neg B$).

A nonmonotonic conditional represents how evidence should affect an (ideally rational) agent's beliefs. For a given conditional, all of the following conditional assertions can hold simultaneously: $C \rightharpoondown B$, $D\&C \rightharpoondown \neg B$, $E\&D\&C \not\rightharpoondown B$ and $E\&D\&C \not\rightharpoondown \neg B$, $F\&E\&D\&C \rightharpoondown B$. So, just as an agent's conditional probability function models the rise and fall of her degrees of belief as new evidence accumulates, an agent's nonmonotonic conditional models the logic of qualitative belief change. The logics of the nonmonotonic conditionals I will investigate are intended to accommodate qualitative defeasible reasoning in the same way that the logic of conditional probabilities accommodates such reasoning when probability values are available. I will treat these conditionals as metalinguistic relations between sentences that satisfy certain axioms, just like conditional probabilities. The role that these conditionals are to play in defeasible reasoning is precisely analogous to the role played by conditional probabilities. Just as the *logic* of an agent's probabilistic inferences is captured by the axioms of probability theory, the *logic* of the uncertain inferences warranted by an agent's nonmonotonic conditional will be captured by axioms for the conditional, axioms which permit the derivation of some conditional assertions from other conditional assertions that the agent maintains. The axioms for nonmonotonic conditionals will play the same kind of dual role played by the axioms for probabilities. They are metalinguistic (semantic) rules that function as a sound (and complete) collection of rules through which some conditional assertions may be derived from others; and they supply constraints that any nonmonotonic conditional must satisfy if it is to yield a consistent assignment of defeasible support for sentences.

As I said earlier, I will explicate a range of different types of nonmonotonic conditionals in this paper, conditionals for which '$C \rightarrow B$' means 'if C, then almost certainly B', conditionals for which '$C \rightarrow B$' means 'if C, then *just possibly* B', and conditionals that correspond to the range of support levels in between. The various types of conditionals will have a number of axioms (i.e. semantic rules) in common, but their axioms will diverge as appropriate for the different support levels to which they correspond. For some of these conditionals '$C \rightarrow B$', and '$C \rightarrow A$' and '$C \not\rightarrow B\&A$' may all hold, and may even be consistent with '$C \rightarrow \neg(B\&A)$'. Some of these conditionals will accommodate the lottery "paradox" – i.e. $E \rightarrow \neg H_1$, $E \rightarrow \neg H_2, \ldots, E \rightarrow \neg H_n$, are consistent with $E \rightarrow H_1 \vee H_2 \vee \cdots \vee H_n$. That is, for each person b, 'b won't win' is *supported* by E, but 'one of these people will win' is also supported by E. For some other conditionals, which correspond roughly to 'if C, then it is just possible that B', the evidence C may even *support* each of a long list of pairwise logically inconsistent sentences.

2. CONDITIONAL-PROBABILISTIC LOGIC

In this section I will explicate a logic for conditional probability functions. Throughout the remainder of the paper conditional probability and its logic will serve as a standard to which the various types of nonmonotonic conditionals and their logics will be compared. Throughout this paper I will restrict attention to an object language L for sentential logic. The language of sentential logic suffices to capture the most salient features of probabilistic logic, and it will also suffice for the explication of the most salient features of nonmonotonic logics. L may contain a finite or a countably infinite set of sentence letters, and contains all sentences constructable from them in the usual way with the logical connectives '$\neg$' and '$\&$' and parentheses; expressions of form '$(A \vee B)$', '$(A \supset B)$', and '$(A \equiv B)$' abbreviate '$\neg(\neg A\&\neg B)$', '$\neg(A\&\neg B)$', and '$(\neg(A\&\neg B)\&\neg(B\&\neg A))$', respectively (and I will often drop parentheses when no ambiguity threatens). 'S_L' denotes the set of sentences of L. Capital letters 'A', 'B', 'C', etc. with or without subscripts are metalinguistic variables representing sentences of the object language. It will sometimes be convenient to employ some standard tautology or contradiction, so I will let '**T**' and '**F**' abbreviate the two sentences of form '$\neg(A\&\neg A)$' and '$(A\&\neg A)$', respectively, with the first sentence letter of L in place of the 'A'.

The *logical consequence* relation is usually defined in terms of the set of possible truth-value assignments to sentences in S_L. The set of possible truth-value assignments may be defined as follows:

DEFINITION 1. $v \in \mathbf{TVA}$ iff $v \subseteq S_L$ such that, for all A and B in S: (where '$v[C]$' just abbreviates '$C \in v$')

(1) $v[\neg A]$ if and only if not $v[A]$
(2) $v[(A\&B)]$ if and only if $v[A]$ and $v[B]$.

A is *logically true*, i.e. '$\models A$', iff for all $v \in \mathbf{TVA}$, $v[A]$;

A is a *logical consequence* of B, i.e. '$B \models A$', iff for every v in $\mathbf{TVA}$, if $v[B]$, then $v[A]$.

Classical probability is usually defined in a way that depends on a pre-defined notion of *logical truth*, and so ultimately depends on $\mathbf{TVA}$. The set of classical probability functions may be specified as follows.

DEFINITION 2. $P \in \mathbf{CLASSPROB}$ iff P is a function from S_L into $[0, 1]$ such that, for all A, B in S:

(1) if $\models A$, then $P[A] = 1$
(2) if $\models \neg(A\&B)$, then $P[A \vee B] = P[A] + P[B]$.

The *conditional probability* of A given B, $P[A \mid B]$, is defined as follows:

$$P[A \mid B] = P[A\&B] \div P[B] \quad \text{if } P[B] \neq 0;$$

$$P[A \mid B] = 1, \quad \text{if } P[B] = 0.$$

$\mathbf{CLASSPROB}$ is the usual set of classical probability functions on sentences. The definition of *conditional probability* slightly extends the usual definition, which leaves $P[A \mid B]$ undefined if $P[B] = 0$.

Conditional probability is not a primitive part of the classical definition of probability, but is defined in terms of the unconditional notion. In 1938 Karl Popper proposed an axiom system for conditional probability that takes *conditional* probabilities as primitive (see the *New appendices* of (Popper 1968)). Popper's idea was to develop a logic for conditional probabilities that is autonomous from deductive logic. The following definition of the *Popper Functions* is equivalent to Poppers axiomatizations.

DEFINITION 3. $P \in$ **POPPERFN** iff P is a function from $S_L \times S_L$ into $[0, 1]$ such that:

(1) for some D and E in S_L, $P[D \mid E] \neq 1$;
and for all A, B, C in S_L,
(2) $P[A \mid A] = 1$
(3) $P[A \mid C\&B] = P[A \mid B\&C]$
(4) $P[B\&A \mid C] = P[A\&B \mid C]$
(5) $P[A \mid B] + P[\neg A \mid B] = 1$ or $P[C \mid B] = 1$
(6) $P[A\&B \mid C] = P[A \mid B\&C] \times P[B \mid C]$.

The unconditional probability of A may then be defined as $P[A \mid \mathbf{T}]$.

Notice that the rules for **POPPERFN** do not presuppose the substitutivity of logically equivalent sentences, nor do they make any other use of **TVA** and its notion of logical consequence. Conditional probability functions are primitive in the same sense that truth-value assignments are primitive for **TVA**. Indeed, this way of defining the Popper Functions shows them to be a generalization of truth-value semantics. For, one may *define* a probabilistic notion of *logical consequence*, $\Rightarrow$, as follows.

DEFINITION 4. $B \Rightarrow A$ iff for all $P \in$ **POPPERFN**, $P[A \mid B] = 1$.

That is, one may *define* a notion of logical consequence purely in terms of the Popper Functions. This relation turns out to be equivalent to the classical notion of logical consequence.

THEOREM 1. $B \Rightarrow A$ *if and only if* $B \models A$.

Proof. The direction from left to right is obvious since clearly there is a Popper Function that only assigns $P[A \mid B] = 1$ where $B \models A$. The other direction takes work. Popper effectively proved the result by showing that, for each conditional probability function that satisfies his version of the rules, the laws of Boolean algebra hold for the sentences in S_L. Field (1977) gives a different kind of proof, and he also extends the theorem to a language for predicate logic with quantifiers. See also (Harper 1975), (Leblanc 1979, 1983), and (van Fraassen 1981) for formal treatment of the Popper Functions. □

The Popper Functions are really a fairly mild extension of the classical probability functions. They may be more easily compared to classical probability when they are defined in a more classical fashion.

DEFINITION 5. $P \in$ **CONDPROB** iff P is a function from $S_L \times S_L$ into $[0, 1]$ such that:

(1) for some E and G in S_L, $P[G \mid E] \neq 1$;
and for all A, B, C, D and S_L,
(2) if $\models C \equiv B$, then $P[A \mid B] = P[A \mid C]$
(3) if $C \models A$, then $P[A \mid C] = 1$
(4) if $C \models \neg(A\&B)$, then either
$P[A \vee B \mid C] = P[A \mid C] + P[B \mid C]$ or $P[D \mid C] = 1$
(5) $P[A\&B \mid C] = P[A \mid B\&C] \times P[B \mid C]$.

THEOREM 2. **POPPERFN = CONDPROB.**

Proof. With help from Theorem 1 just show that each set of rules is derivable from the other. □

The rules of **CONDPROB** are the obvious extensions of the rules of **CLASSPROB** to a logic in which conditional probability is primitive. Clearly, all of the conditional probability functions defined in terms of **CLASSPROB** are members of **CONDPROB**. But notice that rule 5 permits **CONDPROB** to have members in which $P[A\&B \mid C] = P[B \mid C] = 0$ and yet $P[A \mid B\&C]$ is positive and less than 1. The members of **CONDPROB** behave like classical probabilities except that conditionalization on a sentence with probability 0 may provide a non-classical transition to a *probability space at a lower level*. The way that this works will be described precisely in Section 3.

The definition of conditional probability functions in terms of the rules of **POPPERFN** has interesting implications. The usual definition of conditional probability suggests that probabilistic logic *presupposes* a logic based on truth-values to define the notion of logical truth that it employs. But the rules of **POPPERFN** define conditional probability functions as primitive semantic operations, independent of any predefined notion of logical truth, much as the rules of **TVA** define truth-value assignments as primitive semantic predicates. Thus, the logic of probabilistic support need not presuppose a logic of truth-values.[4]

3. THE LOGIC OF NEAR CERTAINTY

In this section I will investigate the relationship between the Popper Functions and a class of conditionals called *Rational Consequence* relations by Lehmann and Magidor (1992). These conditionals turn out to be the probability 1 parts of Popper Functions. Intuitively a conditional assertion '$B \rightarrow A$' for a Rational Consequence relation says that among the possible states of affairs in which B is true, A is almost certainly true. Only under very exceptional circumstances will the support afforded A by B be undermined, i.e. if '$B \rightarrow A$' holds, then '$C\&B \rightarrow A$'

can fail to hold *only if* '$B \rightarrow \neg C$' holds. Thus, when an agent employs a Rational Consequence relation for defeasible reasoning, the near certainly that B affords A (when B is the total relevant evidence) cannot be undermined by additional information C *unless* C was presumed to be almost certainly false on the basis of B alone.

In the first subsection of the present section I will define a set of conditionals directly in terms of the Popper Functions. Then I will define another set of conditionals in terms of qualitative rules that are analogs of the quantitative rules for Popper Functions. In the second subsection I will define the set of Rational Consequence relations in (roughly) the usual way, and show that all three definitions specify the same set of conditionals. Each definition reveals a different facet of Rational Consequence relations.

3.1. *Conditional Probability* 1 *and the* **ER** *Conditionals*

Consider the set of conditionals that behave like the probability 1 parts of conditional probability functions, i.e. the set of conditionals defined as follows:

DEFINITION 6. $\rightarrow \in$ **POPPERFN**[1] <u>iff</u> for some $P \in$ **POPPERFN**, $\rightarrow = \{\langle B, A\rangle \mid P[A \mid B] = 1\}$.

Each conditional, $\rightarrow$, in **POPPERFN**[1] is obtained by selecting a probability function P from **POPPERFN** and writing '$B \rightarrow A$' just when $P[A \mid B] = 1$ (and writing '$B \not\rightarrow A$' otherwise). Can we specify rules (or axioms) directly in terms of conditionals that capture precisely the conditionals in **POPPERFN**[1]? One way to get rules is to read them off of the rules for conditional probabilities, although this method will not *necessarily* generate enough rules to capture **POPPERFN**[1] completely. The next definition does just that. It employs qualitative versions of rules from **POPPERFN** (Definition 3) to define a set, **ER**, of nonmonotonic *Entailment Relations*.

DEFINITION 7. $\rightarrow \in$ **ER** <u>iff</u> $\rightarrow \subseteq S_L \times S_L$ such that, for all A, B, C in S_L, $\rightarrow$ satisfies the following rules:

(1) for some D and E in S_L, $E \not\rightarrow D$
(2) $A \rightarrow A$
(3) if $C\&B \rightarrow A$, then $B\&C \rightarrow A$
(4.1) if $C \rightarrow B\&A$, then $C \rightarrow A\&B$
(4.2) if $C \rightarrow \neg(B\&A)$, then $C \rightarrow \neg(A\&B)$
(5.1) if $B \rightarrow \neg\neg A$, then $B \rightarrow A$

(5.2) if $B \to A$ and $B \to \neg A$, then $B \to C$
(6.1) $C \to B$ and $C\&B \to A$ iff $C \to B\&A$
(6.2) $C \to \neg B$ or $C\&B \to \neg A$ iff $C \to \neg(B\&A)$.

Rules for nonmonotonic conditionals are commonly represented in the literature as deduction rules, but I think they should be understood as semantic rules, akin to the rules for probability functions. However, in this paper I will often write the semantic rules for conditionals in a form to look like deduction rules in order to facilitate comparison with the deduction rule notation employed elsewhere in the literature. Rules 2–6 together with a reasonable substitute for rule 1 take the following form:

$$(1)\ \frac{}{\mathbf{T} \not\to \mathbf{F}} \qquad (2)\ \frac{}{A \to A} \qquad (3)\ \frac{C\&B \to A}{B\&C \to A}$$

$$(4.1)\ \frac{C \to B\&A}{C \to A\&B} \qquad (4.2)\ \frac{C \to \neg(B\&A)}{C \to \neg(A\&B)}$$

$$(5.1)\ \frac{B \to \neg\neg A}{B \to A} \qquad (5.2)\ \frac{B \to A,\ B \to \neg A}{B \to C}$$

$$(6.1)\ \frac{C \to B,\ C\&B \to A}{C \to B\&A} \qquad \frac{C \to B\&A}{C \to B}$$

$$\frac{C \to B\&A}{C\&B \to A}$$

$$(6.2)\ \frac{C \to \neg B}{C \to \neg(B\&A)} \qquad \frac{C\&B \to \neg A}{C \to \neg(B\&A)}$$

$$\frac{C \not\to \neg B,\ C\&B \not\to \neg A}{C \not\to \neg(B\&A)}.$$

The rules for **ER** come directly from the correspondingly numbered rules for **POPPERFN** by taking the values of the probabilities to be 1 (or, 0 in some cases). The rules of **ER** do not presuppose that logically equivalent sentences may be substituted. They are completely autonomous with regard to **TVA** and its logical consequence relation. Indeed, **ER** has its own notion of logical consequence.

DEFINITION 8. $B \Rightarrow A$ (A is an **ER**-logical consequence of B) iff for all $\to \in$ **ER**, $B \to A$.

The **ER** notion of logical consequence is equivalent to the classical notion.

THEOREM 3. *$B \Rightarrow A$ if and only if $B \models A$.*

Proof. It's easy to check that $\models$ is a member of **ER** – it satisfies the rules. So, if $B \Rightarrow A$, then $B \models A$. The proof in the other direction takes more work. Suppose $B \not\Rightarrow A$. Then for some $\rightarrow \in$ **ER**, $B \not\rightarrow A$. For this $\rightarrow$ one can derive that $B\&\neg A \not\rightarrow A$. The remainder of the proof is like a Henkin proof. Order all of the sentences of S_{L}. If for the first sentence, C, $B\&\neg A \rightarrow C$ or $B\&\neg A \rightarrow \neg C$, move on to the next sentence. Otherwise, $(B\&\neg A)\&C \not\rightarrow A$. It can then be shown that there is a conditional $\rightarrow'$ such that for all sentences D, $(B\&\neg A) \rightarrow' D$ just in case $(B\&\neg A)\&C \rightarrow D$. Now continue the construction using $\rightarrow'$. The set v of all sentences D such that for some conditional $\rightarrow''\cdots'$ in the construction, $B\&\neg A \rightarrow''\cdots' D$, is a truth-value assignment that makes $(B\&\neg A)$ true. □

Many of the rules of **ER** are close relatives of more familiar rules for Rational Consequence relations presented in Lehmann and Magidor (1992). Indeed, **ER** will turn out to be the set of Rational Consequence relations, and in turn they are just the probability 1 parts of the Popper Functions. The rules of **ER** define nonmonotonic conditionals as primitive semantic relations on sentences in precisely the way that the rules of **TVA** define the truth-value assignments as primitive semantic predicates. Thus, like the logic of probabilistic support rendered by the Popper Functions, the logic of defeasible support need not essentially depend on the theory of truth. In the next subsection I will characterize the Rational Consequence relations in a more usual way. The rules in that characterization are related to the rules of **ER** in the same way that the rules of **CONDPROB** are related to those of **POPPERFN**.

3.2. *The Rational Consequence Relations*

I will first define a basic set of conditionals that I call **O**. All of the conditionals investigated in this paper will satisfy the rules of **O**.

DEFINITION 9. $\rightarrow \in$ **O** iff $\rightarrow \subseteq S_{\mathrm{L}} \times S_{\mathrm{L}}$ such that, for all A, B, C in S, $\rightarrow$ satisfies the following rules:

(1) $\mathbf{T} \not\rightarrow \mathbf{F}$	(Nondegeneracy)
(2) if $\models B \equiv C$ and $B \rightarrow A$, then $C \rightarrow A$	(Left Logical Equivalence)
(3) if $C \rightarrow B$ and $\models B \supset A$, then $C \rightarrow A$	(Right Weakening)
(4) $A \rightarrow A$	(Reflexivity)
(5) if $C\&B \rightarrow A$ and $C\&\neg B \rightarrow A$, then $C \rightarrow A$	(Weak Or)

(6) if $C \rightarrow B\&A$, then $C\&B \rightarrow A$ (Very Cautious Monotonicity)

(7) if $B \rightarrow A$, then $B \rightarrow B\&A$ (Weak And)

(8) if $C \rightarrow \neg C$ and $B \rightarrow A$, then $B \rightarrow \neg C\&A$ (Conjunctive Certainty).

In deduction rule form rules 1–8 are as follows:

$$(1)\quad \frac{}{\mathbf{T} \not\rightarrow \mathbf{F}} \qquad (2)\quad \frac{\models B \equiv C,\ B \rightarrow A}{C \rightarrow A}$$

$$(3)\quad \frac{C \rightarrow B,\ B \models A}{C \rightarrow A} \qquad (4)\quad \frac{}{A \rightarrow A}$$

$$(5)\quad \frac{C\&B \rightarrow A,\ C\&\neg B \rightarrow A}{C \rightarrow A} \qquad (6)\quad \frac{C \rightarrow B\&A}{C\&B \rightarrow A}$$

$$(7)\quad \frac{B \rightarrow A}{B \rightarrow B\&A} \qquad (8)\quad \frac{C \rightarrow \neg C,\ B \rightarrow A}{B \rightarrow \neg C\&A}.$$

The conditionals in **O** are *probabilistically sound*. That is, any conditional that behaves like the probability-greater-than-r part of a conditional probability function must obey the rules of **O**. More precisely, if we pick a number r between 0 and 1 and read each conditional assertion of form '$X \rightarrow Y$' as $P[Y \mid X] \geqslant r$, then each of rules 1–8 are theorems about these conditionals. Rule 5, for example, is probabilistically sound because, for any r between 0 and 1 and any P in **POPPERFN**, if $P[A \mid C\&B] \geqslant r$ and $P[A \mid \neg C\&B] \geqslant r$, then $P[A \mid B] \geqslant r$. Not all probabilistically sound rules are derivable from the rules of **O**. In Section 4 they will be supplemented with another sound rule.

Ernest Adams (1966, 1975) was first to work out a precise logical connection between nonmonotonic conditionals and conditional probabilities. The connection Adams discovered between nonmonotonic conditionals and conditional probabilities involves classical conditional probabilities, not the Popper Functions. He recognized that when classical conditional probabilities are very nearly 1 they exhibit the logic of nonmonotonicity that indicative conditionals should have. Adams characterizes the connection between the logic of conditionals and probabilities through the notion of *p-entailment*, as follows:

DEFINITION 10. $\{(B_1 \rightarrow A_1), \ldots, (B_n \rightarrow A_n)\}$ *p-entails* $(B \rightarrow A)$ iff for each $\varepsilon > 0$ there is a $\delta > 0$ such that for each P in **CLASSPROB**, if $P[A_i \mid B_i] \geqslant 1 - \delta$ for each i $(0 \leqslant i \leqslant n)$, then $P[A \mid B] > 1 - \varepsilon$.

Adams developed a set of inference rules for deriving conditional assertions from sets of conditional assertions and showed his rules to provide a sound and complete characterization of p-entailment for finite languages. (See (Pearl 1988) for a nice treatment of Adams' system.) Kraus, Lehmann, and Magidor (1990) call the conditionals that satisfy Adams' inference rules the *Preferential Consequence* relations. Adams' rules are equivalent to the rules of **O** together with rule 9 in the following definition. I will call the set of conditionals that satisfy these rules **P**.

DEFINITION 11. $\rightarrow \in P$ iff $\rightarrow$ satisfies the rules of **O** and the following rule, for all A, B, C in S_{L}:

(9) if $C \rightarrow B$ and $C \rightarrow A$, then $C \rightarrow B\&A$: i.e.

$$\frac{C \rightarrow B,\ C \rightarrow A}{C \rightarrow B\&A} \quad \text{(And)}$$

Rules 7 and 8 of **O** are derivable from rules 1–6 together with rule 9. However, rule 9 is not derivable from **O**'s rules, so **P** is a proper subset of **O**. **P** may employ the following rules rather than 5 and 6:

$$\frac{B \rightarrow A,\ C \rightarrow A}{B \vee C \rightarrow A} \quad \text{(Or)}$$

$$\frac{C \rightarrow A,\ C \rightarrow B}{C\&B \rightarrow A} \quad \text{(Cautious Monotonicity)}$$

Both of these rules are derivable from rules 1–6 and 9. (Or), (Cautious Monotonicity), and (And) are not probabilistically sound rules in the sense specified above, but they are *sound for probabilities of* 1 for Popper Functions. E.g., the rule (And) is sound for probabilities of 1 because, for all P in **POPPERFN**, if $P[A \mid C] = 1$ and $P[B \mid C] = 1$, then $P[A\&B \mid C] = 1$.

Kraus, Lehmann and Magidor (1990) have developed a possible worlds semantics for the conditionals in **P**. Lehmann and Magidor (1992) extend this semantics to a more restricted set of conditionals that they call the *Rational Consequence* relations. Rational relations are defined by adding a stronger monotonicity rule to the rules of **P**. The set of Rational Consequence relations, **R**, may be defined as follows:

DEFINITION 12. $\rightarrow \in \mathbf{R}$ iff $\rightarrow$ satisfies the rules of **P** and the following rule, for all A, B, C in S:

(10) if $C \rightarrow A$ and $C \not\rightarrow \neg B$, then $C\&B \rightarrow A$: i.e.

$$\frac{C \rightarrow A,\ C \not\rightarrow \neg B}{C\&B \rightarrow A} \quad \text{(Rational Monotonicity)}$$

It is routine to derive each rule of **ER** from those of **R**; and the converse is straightforward with the aid of Theorem 3. So, **R** and **ER** are the same sets of conditionals.

THEOREM 4. **R** = **ER**.

Given the similarity between rules of **ER** and those of **POPPERFN** the next theorem is not surprising.

THEOREM 5. **R** = **POPPERFN**[1].

Proof. The hard part of this theorem is essentially proved by Lehmann and Magidor in an appendix to (1992). They show that for each $\rightarrow$ in **R** there is a probability function P on the non-standard real numbers (i.e. the reals with infinitesimals) such that $B \rightarrow A$ if and only if $P[A\&B] \div P[B]$ is infinitesimally close to 1. Define a conditional probability function as follows: $P[A \mid B]$ = the nearest real number to $P[A\&B] \div P[B]$ (i.e. the reals associated with its Dedekind cut), or 1 when $P[B] = 0$. These functions satisfy the rules for **POPPERFN**. Thus, **R** $\subseteq$ **POPPERFN**[1]. And clearly the rules of **R** are satisfied by all probability 1 parts of the *Popper Functions*, so **POPPERFN**[1] $\subseteq$ **R**. □

Thus, the probability 1 parts of Popper Functions are the Rational Consequence relations. And, like the Popper Functions, the Rational relations can be characterized in a way that does not rely on a predefined notion of *logical consequence*. In the next subsection I will briefly characterize some central features of the logics of the *Preferential* and *Rational* relations by describing the restricted versions of monotonicity, transitivity, and contraposition that they satisfy.

3.3. *Some Characteristic Rules of* **O**, **P**, *and* **R**

Monotonic conditionals such as the classical *logical consequence* relation exhibit three properties that fail for nonmonotonic conditionals – monotonicity, transitivity, and contraposition, respectively:

$$\frac{C \rightarrow A}{C\&B \rightarrow A} \qquad \frac{C \rightarrow B,\ B \rightarrow A}{C \rightarrow A} \qquad \frac{C\&B \rightarrow A}{C\&\neg A \rightarrow \neg B}$$

The conditionals in **O**, **P**, and **R** only satisfy weakened versions of these properties.

Regarding monotonicity, **O** has only rule 6 of Definition 9 (Very Cautious Monotonicity). Conditionals in **P** satisfy the first of the following two rules. Conditionals in **R** satisfy both of these rules.

$$\frac{C \rightarrow A,\ C \rightarrow B}{C\&B \rightarrow A} \qquad \text{(Cautious Monotonicity)}$$

$$\frac{C \rightarrow A,\ C \not\rightarrow \neg B}{C\&B \rightarrow A} \quad \text{(Rational Monotonicity)}$$

The conditionals in **O** generally satisfy only extremely weak versions of transitivity, the rules 2 and 3 of **O**. Conditionals in **P** satisfy two moderately stronger transitivity rules, the first two of the following trio. Conditionals in **R** also satisfy the third, somewhat stronger rule.

$$\frac{C \rightarrow B,\ C\&B \rightarrow A}{C \rightarrow A} \qquad \frac{C \rightarrow B,\ B \rightarrow A,\ B \rightarrow C}{C \rightarrow A}$$

$$\frac{C \rightarrow B,\ B \rightarrow A,\ B \not\rightarrow \neg C}{C \rightarrow A}$$

Only an extremely weak version of contraposition holds among all conditionals in **O**, while the weak versions of contraposition for **P** and **R** follow a pattern that may begin to look familiar. Contraposition rules for members of **O**, **P**, and **R**, respectively are these:

$$\frac{C\&B \rightarrow A,\ C \models \neg A}{C\&\neg A \rightarrow \neg B} \qquad \frac{C\&B \rightarrow A,\ C \rightarrow \neg A}{C\&\neg A \rightarrow \neg B}$$

$$\frac{C\&B \rightarrow A,\ C \not\rightarrow A}{C\&\neg A \rightarrow \neg B}$$

A few additional observation are in order before moving on. All of the conditionals studied in this paper satisfy the rules of **O**, so the above rules for members of **O** will apply throughout. It is also worth noting that for any $\rightarrow$ in **O**, '$C \rightarrow \neg C$' holds just in case '$C \rightarrow \mathbf{F}$' holds, so when a sentence C satisfies '$C \rightarrow \neg C$' we will call C *inconsistent in* $\rightarrow$. If C is inconsistent in $\rightarrow$, C will be irrelevant to other conditional assertions in the sense expressed in rule 8 of **O**'s definition and the following derived rules:

$$\frac{C \rightarrow \neg C,\ B \rightarrow A}{B\&\neg C \rightarrow A} \qquad \frac{B \models \neg C,\ C \rightarrow \neg C,\ B \vee C \rightarrow A}{B \rightarrow A}$$

$$\frac{A \models \neg C,\ C \rightarrow \neg C,\ B \rightarrow A \vee C}{B \rightarrow A}$$

The next two rules are also derivable in **O**. The first is equivalent to (Weak Or) given the other rules:

$$\frac{\models \neg(C\&D),\ C \rightarrow A,\ D \rightarrow B}{C \vee D \rightarrow A \vee B} \qquad \frac{B \rightarrow A}{C \vee B \rightarrow B \supset A}.$$

3.4. *Orderings and Rankings*

Every Conditional in **R** imposes a ranking on sentences that completely characterizes the conditional. A sentence of lower rank is a much more weighty possibility than any sentence ranked above it, in a sense to be made more precise in this subsection. Kraus, Lehmann and Magidor (1990), and Lehmann and Magidor (1992) have thoroughly investigated the orderings and rankings imposed by conditionals in **P** and **R**. However, they did not investigate **O**. The orderings on sentences imposed by conditionals in **O** will be important to the treatment of the conditionals in Section 4, so I will treat these orderings in some detail here.

DEFINITION 13. For each conditional, $\rightarrow$, in **O** define a relation $\gg$ on sentences in S_{L} as follows:

$A \gg B$ <u>iff</u> $A \not\rightarrow \mathbf{F}$, and for all D, C in S_{L}, if $D \vee A \rightarrow C$, then $D \vee A \rightarrow \neg B\&C$ – i.e. <u>iff</u> $A \not\rightarrow \mathbf{F}$, and for all C, D in S_{L}, the following rule holds for $\rightarrow$:

$$\frac{D \vee A \rightarrow C}{D \vee A \rightarrow \neg B\&C}$$

Also define: $A \approx B$ <u>iff</u> not $A \gg B$ and not $B \gg A$;

$A \underline{\gg} B$ <u>iff</u> $A \gg B$ or $A \approx B$.

On a possible worlds reading of conditionals the conditional assertion '$A \rightarrow C$' says that among the worlds in which A is true, the subset of worlds in which C is true is weighty enough for C to be provisionally accepted. Then, '$A \gg B$' says that among *any* set of possible worlds that contains all of the worlds in which A holds, the subset of worlds in which B holds is so insignificant that support for any proposition C depends only on the weight of C worlds in which B is false.

If $\rightarrow$ is in **O**, its ordering relation $\gg$ is asymmetric and transitive, i.e. is a *strict partial ordering*.

THEOREM 6. *For each* $\rightarrow \in$ **O** *its relation* $\gg$ *is asymmetric and transitive and* $\approx$ *is symmetric and reflexive.*

The following relationship also hold:

$A \approx B$ if and only if $A \underline{\gg} B$ and $B \underline{\gg} A$;
$A \gg B$ if and only if $A \underline{\gg} B$ and not $B \underline{\gg} A$;
exactly one of $A \gg B$, $B \gg A$, or $A \approx B$ must hold;
$A \underline{\gg} B$ or $B \underline{\gg} A$, i.e. $\underline{\gg}$ is connected.

Also, $T \gg \mathbf{F}$; *if* $A \gg B$, *then* $A \vee C \gtrsim B \vee C$ *and* $A \vee C \gg B$; *if* $B \models A$, *then* $A \gtrsim B$.

Proof. Clearly $\gg$ is asymmetric. For if $A \gg B \gg A$, then it follows from $A \vee B \rightarrow A \vee B$, that $A \vee B \rightarrow \neg A \& \neg B \& (A \vee B)$, so we have $A \vee B \rightarrow \mathbf{F}$. But then $A \vee B \rightarrow A \& \mathbf{F}$, so $A \rightarrow \mathbf{F}$, which contradicts $A \gg B$. To see that $\gg$ is transitive, suppose that $A \gg B \gg C$, and suppose that for some D and E, $D \vee A \rightarrow E$. Then $D \vee A \rightarrow \neg B \& E$, so $(D \vee A \vee B) \& \neg B \rightarrow E$, so by a derived rule for $\mathbf{O}$, $(D \vee A \vee B) \rightarrow \neg B \supset E$. Then, since $A \gg B$, $(D \vee A \vee B) \rightarrow \neg B \& (\neg B \supset E)$. Thus, $D \vee A \vee B \rightarrow \neg B \& E$. Since $B \gg C$, $D \vee A \vee B \rightarrow \neg C \& \neg B \& E$. It follows that $D \vee A \vee B \rightarrow (D \vee A \vee B) \& \neg C \& \neg B \& E$, so $D \vee A \vee B \rightarrow (D \vee A) \& \neg C \& E$ by (Right Weakening). Then by (Very Cautious Monotonicity) and (Left Logical Equivalence) we have $D \vee A \rightarrow \neg C \& E$. This derivation holds for arbitrary D and E. Thus, $A \gg C$. The other properties stated in the theorem can be easily derived. □

All it would take for $\gtrsim$ to become a weak order (i.e. transitive and connected order) is for $\approx$ to be transitive. But $\approx$ may not be transitive for some members of $\mathbf{O}$.

Kraus, Lehmann and Magidor in (1990) and Lehmann and Magidor in (1992) describe strict partial orders on sentences induced by conditionals in $\mathbf{P}$. The strict partial orders, $\gg$, described above are essentially the orderings they investigate. The additional rule (And) that $\mathbf{P}$ imposes on $\rightarrow$ provides a simpler characterization of these ordering for members of $\mathbf{P}$, as the next theorem shows.

THEOREM 7. *For* $\rightarrow \in \mathbf{P}$: $A \gg B$ *iff* $A \not\rightarrow \mathbf{F}$ *and* $A \vee B \rightarrow \neg B$.

Proof. For any $\rightarrow$ in $\mathbf{O}$, if $A \gg B$, then $A \not\rightarrow \mathbf{F}$; and from $A \vee B \rightarrow A \vee B$, we have $A \vee B \rightarrow \neg B \& (A \vee B)$, so $A \vee B \rightarrow \neg B$. The proof of the other direction needs (And). Suppose $A \vee B \rightarrow \neg B$ and $A \not\rightarrow \mathbf{F}$. Then in $\mathbf{O}$ it follows that for any sentence D, $D \vee A \vee B \rightarrow \neg B$ (from (Weak Or) since $(D \vee A \vee B) \& (A \vee B) \rightarrow \neg B$ and $(D \vee A \vee B) \& \neg(A \vee B) \rightarrow \neg B$). Then $(D \vee A \vee B) \rightarrow (D \vee A \vee B) \& \neg B$, so $(D \vee A \vee B) \rightarrow (D \vee A) \& \neg B$, so $(D \vee A \vee B) \& (D \vee A) \rightarrow \neg B$, then $D \vee A \rightarrow \neg B$. This holds for every sentence D. Now suppose $D \vee A \rightarrow E$ for some D and E. Then from (And) we have $D \vee A \rightarrow \neg B \& E$. Thus, $A \gg B$. □

Thus, if a conditional $\rightarrow$ is in $\mathbf{P}$, then the strict partial order that $\rightarrow$ already imposes on sentences due to its membership in $\mathbf{O}$ is completely determined by the conditional assertions of form '$A \vee B \rightarrow \neg B$' that hold. Conditionals in $\mathbf{R}$ belong to $\mathbf{P}$, and so share this property. But, in addition, for members of $\mathbf{R}$ the relations $\approx$ turns out to be transitive, and this suffices to make $\gg$ a weak order (transitive and connected) which ranks the sentences of S_{L}.

THEOREM 8. *For* $\rightarrow\ \in \mathbf{R}$, $\geqslant\!\!\!\gg$ *is a weak order* (*i.e. transitive and connected*).

Proof. We need only show that $\approx$ is transitive; then, connectedness follows easily. Suppose that $A \approx B \approx C$ and not $A \approx C$. Then $C \gg A$ or $A \gg C$. We need only consider $C \gg A$ (the other case is similar). Since $\rightarrow$ is in $\mathbf{R} \subseteq \mathbf{P}$. Theorem 6 yields $A \vee B \vee C \gg A$; Theorem 7 implies $A \vee B \vee C \rightarrow \neg A$. Now, if $A \vee B \vee C \rightarrow \neg A \& \neg B$, then $A \vee B \vee C \rightarrow (B \vee C) \& \neg A \& \neg B$, so $(A \vee B \vee C) \& (B \vee C) \rightarrow \neg B$, and we have $B \vee C \rightarrow \neg B$ which by Theorem 7 violates $B \approx C$. So instead we must have $A \vee B \vee C \not\rightarrow \neg A \& \neg B$. Then, $A \vee B \vee C \not\rightarrow \neg(A \vee B)$. This together with $A \vee B \vee C \rightarrow \neg A$ and (Rational Monotonicity) yields $(A \vee B \vee C) \& (A \vee B) \rightarrow \neg A$; so $A \vee B \rightarrow \neg A$. Then, $B \gg A$ (Theorem 7), which contradicts $A \approx B$. Thus, $\approx$ is transitive. □

The weak order that a conditional in $\mathbf{R}$ imposes on sentences ranks all sentences in S_{L}.

DEFINITION 14. For each $\rightarrow\ \in \mathbf{R}$ define the rank function for $\rightarrow$ as follows:

$$\begin{aligned}
\text{rank}[A] &= 1 \ \underline{\text{iff}}\ \mathbf{T} \approx A;\\
\text{rank}[A] &= i+1 \ \underline{\text{iff}}\ A \not\rightarrow \mathbf{F},\ \text{rank}[A] \not\leqslant i,\ \text{and for all } B,\\
&\qquad \text{rank}[B] \leqslant i \text{ or } B \rightarrow \mathbf{F} \text{ or } A \geqslant\!\!\!\gg B.\\
\text{rank}[A] &= \infty \ \underline{\text{iff}}\ A \rightarrow \mathbf{F}.
\end{aligned}$$

Notice that rank order runs in the opposite direction of the weak ordering that generates it – i.e. the weightier possibilities have the lower rank. Theorem 5 showed that each probability function in **POPPERFN** "contains" a conditional in $\mathbf{R}$. So conditional probability functions impose precisely the same sort of ranking on sentences as conditionals in $\mathbf{R}$ (i.e. define '$A \gg B$' as '$P[\mathbf{F} \mid A] < 1$, and for all D, C, if $P[C \mid D \vee A] = 1$, then $P[\neg B \& C \mid D \vee A] = 1$'; and replace occurrences of '$B \rightarrow A$' in theorems of this subsection with '$P[A \mid B] = 1$'). The theorem following the next definition will show the fundamental role played by rank for conditionals in $\mathbf{R}$ and for Popper Functions.

DEFINITION 15. For a finite language, L a *state description* of L is any conjunction of literals (i.e. sentence letters of L and their negations) that contain each sentence letter of L or its negation (but not both). Let SD be the set of all state descriptions of L. For any countable language, L (possibly infinite) and any finite set of sentences $\{A, \ldots, C\}$ of L, define SD$\{A, \ldots, C\}$ as the set of state descriptions for the finite language consisting of the sentence letters occurring in the sentences of the set $\{A, \ldots, C\}$.

THEOREM 9. *For* $\to\, \in \mathbf{R}$, $B \to A$ *iff either* $B \to \mathbf{F}$, *or for all* $C \in$ SD$\{A, B\}$ *such that* rank$[C] =$ rank$[B]$, *if* $C \models B$, *then* $C \models A$.

Proof. (1) Suppose the right-hand side of the theorem is satisfied. If $B \to \mathbf{F}$, then $B \to A$. So, suppose $B \not\to \mathbf{F}$. It will suffice to prove the theorem for sentences B and A that are disjunctions of members of SD$\{A, B\}$ (since they will be logically equivalent to sentences in this form). Let $\{B_1, \ldots, B_n\}$ and $\{A_1, \ldots, A_n\}$ be the subsets of SD$\{A, B\}$ that are disjuncts of B and not of A, and disjuncts of A but not of B, respectively. Let $\{C_1, \ldots, C_k\}$ be members of SD$\{A, B\}$ shared by A and B. Then, $B \to A$ just in case $B_1 \vee \cdots \vee B_n \vee C_1 \vee \cdots \vee C_k \to C_1 \vee \cdots \vee C_k$ (since $B \to A$ iff $B \to B\&A$). Notice that $\{C_1, \ldots, C_k\}$ is not empty unless $B \to \mathbf{F}$. Let $B \gg C_i$ for just the C_i such that $i > h$; and let $B \gg B_j$ for just the $j > g$. Then $B_1 \vee \cdots \vee B_n \vee C_1 \vee \cdots \vee C_k \to C_1 \vee \cdots \vee C_k$ holds just in case $B_1 \vee \cdots \vee B_g \vee C_1 \vee \cdots \vee C_k \to (C_1 \vee \cdots \vee C_h)$. So, $B \to A$ iff $B_1 \vee \cdots \vee B_g \vee C_1 \vee \cdots \vee C_h \to (C_1 \vee \cdots \vee C_h)$, where all the state descriptions involved on the right of the 'iff' are the same rank as B. Now, since we are assuming that the right-hand side of the theorem holds, $\{B_1, \ldots, B_g\}$ must be empty. Therefore, $B \to A$ holds, since clearly $C_1 \vee \cdots \vee C_h \to (C_1 \vee \cdots \vee C_h)$ holds.

(2) For the other direction, suppose $B \to A$. Let $D \in$ SD$\{A, B\}$ be such that $D \models B$ and not $D \models A$. Then, since D is a state description, $D \models \neg A$. So, $A \models \neg D$. Then $B \to \neg D\&A$, thus $B \to \neg D$. But $\models B \equiv (B \vee D)$, so $B \vee D \to \neg D$. Therefore, by Theorem 7, $B \gg D$ or $B \to \mathbf{F}$. □

Theorem 9 applies to both finite and countably infinite languages. But consider for a moment languages with only a finite set of sentence letters. This will provide a simpler picture of the theorem's implications. Imagine a truth table for finite language L with each sentence letter represented across the top. Each line of the truth table makes exactly one state description of L true. Each sentence of L may be thought of as representing the set of truth table lines that make it true. We can generate conditionals in $\mathbf{R}$ by putting *rankings* on truth table lines in the following way. Label some or all of the lines of the truth table with a '1' (for rank 1). For the unlabeled lines that remain (if any), label some (or all) with '2', or else label *all* remaining lines with '∞'. Continue in this way until each line is either labeled with some number or with '∞'.

Define the rank of logical contradictions to be ∞, too. Define the rank of any sentence in S_L as that of the lowest numbered line that makes it true. This ranked truth table yields a relation $\rightarrow$ in **R**, as follows: define $B \rightarrow A$ to hold *just in case* either B is of rank ∞, or else every truth table line that has the same rank as B and makes B true also makes A true. It is easy to check that $\rightarrow$ is a member of **R** (it satisfies the rules of **R**). Theorem 9 also implies the converse, that every conditional in **R** can be generated in this way.

We may generate conditional probability functions in **POPPERFN** from ranked truth tables as follows. Label every truth table line within the same rank (except for rank ∞) with a second number between 0 and 1, its probability, so that the sum of these numbers adds to 1 when summed for that rank; and label lines of rank ∞ with "probability" 0. Define $P[A \mid B]$ = "the sum of the probabilities of truth table lines that have the same rank as B and make $A\&B$ true, divided by the sum of the probabilities of the rank B lines that make B true." It is easy to check that all such functions P are Popper Functions, and all Popper Functions reduce to such ranked truth tables. Thus, Popper Functions are basically just a nested hierarchy of classical probability functions on ranked interpretations of a formal language.

4. THE LOGICS OF LEVELS OF PROBABILISTIC SUPPORT

The precise relationship between the conditionals in **R** and the Popper Functions was given by Theorem 5, $\mathbf{R} = \mathbf{POPPERFN}[1]$. A natural extension of the definition of $\mathbf{POPPERFN}[1]$ to conditionals corresponding to probabilistic support at a level less than 1 might plausibly go like this:

$$\rightarrow \in \mathbf{POPPERFN}[p] \ \underline{\text{iff}} \text{ for some } P \in \mathbf{POPPERFN},$$
$$\rightarrow = \{\langle B, A\rangle \mid P[A \mid B] \geqslant p\}.$$

This definition turns out not to be the most useful way to construct classes of conditionals from Popper Functions. Identical sets of qualitative rules will hold for conditionals corresponding to a range of values of support level p. So, a slightly more general extension of $\mathbf{POPPERFN}[1]$ will provide the most natural way to classify the conditionals that correspond to various support levels.

DEFINITION 16. For $0 \leqslant p < q \leqslant 1$, $\rightarrow \in \mathbf{POPPERFN}[p,q]$ <u>iff</u> there is a $P \in \mathbf{POPPERFN}$ such that for each rank i of P there is a real number r_i, $p < r_i \leqslant q$, such that

$$\rightarrow = \{\langle B, A\rangle \mid \text{rank}_P[B] = \infty,$$

$$\text{or } \text{rank}_P[B] = i \text{ and } P[A \mid B] \geqslant r_i\}.$$

Also, define $\mathbf{POPPERFN}[1, 1] = \mathbf{POPPERFN}[1]$.

Clearly the conditionals in $\mathbf{POPPERFN}[0, 1]$ satisfy the rules of **O**. In the next subsection I will identify an additional rule that all members of $\mathbf{POPPERFN}[0, 1]$ satisfy. Then I will develop additional qualitative rules that are sound for classes of conditionals that correspond to narrower intervals of support levels, rules for the conditionals in $\mathbf{POPPERFN}$ $[1/(n + 1), 1/n]$ and in $\mathbf{POPPERFN}[(n - 1)/n, n/(n + 1)]$, for each $n \geqslant 2$. In subsequent subsections I will show how to supplement the qualitative rules to yield a complete characterization of the support level conditionals for each such interval.

4.1. *The System* **Q** *and the Level Specific Rules*

For each Popper Function P, if $P[A \mid C] \geqslant r$, then either $P[A \mid B\&C] \geqslant r$ or $P[A \mid \neg B\&C] \geqslant r$. This suggests a probabilistically sound rule that is not derivable from the rules for **O**. The rule and its associated set of conditionals (the *Quasi-Probabilistic* consequence relations, **Q**) are specified in the next definition.

DEFINITION 17. $\rightarrow\, \in \mathbf{Q}$ <u>iff</u> $\rightarrow\, \in \mathbf{O}$, and for all A, B, C in S_{L}, $\rightarrow$ satisfies the following rule:

(11) if $C \rightarrow A$, then $C\&B \rightarrow A$ or $C\&\neg B \rightarrow A$: i.e.,

$$\frac{C\&B \not\rightarrow A,\ C\&\neg B \not\rightarrow A}{C \not\rightarrow A} \qquad \text{(Negation Rationality)}$$

The rule (Negation Rationality) is a derived rule of **R**. All of **Q**'s rules are rules of **R**, but (Negation Rationality) is weaker than (Rational Monotonicity), so $\mathbf{R} \subset \mathbf{Q}$. How is **Q** related to the conditionals in $\mathbf{POPPERFN}[p, q]$? The next definition and theorem begin to provide an answer to this question.

DEFINITION 18. Define $\mathbf{Q}[0, 1] = \mathbf{Q}$, and $\mathbf{Q}[1, 1] = \mathbf{R}$. For all integers $n \geqslant 2$, define the following sets of conditionals as those in **Q** that satisfy the specified rule:

$$\mathbf{Q}[1/(n+1), 1]: \quad \frac{C\rightarrow B_1,\ C\rightarrow B_2,\ \ldots,\ C\rightarrow B_{n+1},\ C\models\neg(B_i\&B_j)\ \text{(for each } i\neq j)}{C\rightarrow\neg C}$$

$$\mathbf{Q}[(n-1)/n, 1]: \quad \frac{C\rightarrow\neg B_1,\ C\rightarrow\neg B_2,\ \ldots,\ C\rightarrow\neg B_n,\ C\models(B_1\vee B_2\vee\cdots\vee B_n)}{C\rightarrow\neg C}$$

$\mathbf{Q}[0, n/(n+1)]$:
$$\frac{C \not\rightarrow \neg B_1,\ C \not\rightarrow \neg B_2,\ \ldots,\ C \not\rightarrow \neg B_n,\ C \models \neg(B_i \& B_j)\ \text{(for each } i \neq j,\ i,j = 1,\ldots,n+1)}{C \rightarrow \neg B_{n+1}}$$

$\mathbf{Q}[0, 1/n]$:
$$\frac{C \not\rightarrow B_1,\ C \not\rightarrow B_2,\ \ldots,\ C \not\rightarrow B_{n-1},\ C \models (B_1 \vee B_2 \vee \cdots \vee B_n)}{C \rightarrow B_n}$$

$\mathbf{Q}[p, q] = \mathbf{Q}[p, 1] \cap \mathbf{Q}[0, q]$, for $0 \leqslant p < q \leqslant 1$, with p and q in the set

$$\{\ldots, 1/(n+1),\ 1/n, \ldots, 1/3, 1/2, 2/3, \ldots, (n-1)/n,\\ n/(n+1), \ldots\} \cup \{0, 1\}.$$

Of particular interest are $\mathbf{Q}[(n-1)/n, n/(n+1)] = \mathbf{Q}[(n-1)/n, 1] \cap \mathbf{Q}[0, n/n+1]$ and $\mathbf{Q}[1/(n+1), 1/n] = \mathbf{Q}[1/n+1, 1] \cap \mathbf{Q}[0, 1/n]$. Such intervals partition $\mathbf{Q}$ into the most homogeneous classes of conditionals consistent with the rules of Definition 18. For p and q in any of the fractional units specified in the definition, the rules for $\mathbf{Q}[p, q]$ are just the rules for $\mathbf{Q}$ together with the rule for $\mathbf{Q}[0, q]$ and the rule for $\mathbf{Q}[p, 1]$. Henceforth in expression of form '$\mathbf{Q}[p, q]$' and '$\mathbf{POPPERFN}[p, q]$', p and q are assumed either to both equal 1 or to be in the set $\{\ldots, 1/n, \ldots, 1/3, 1/2, 2/3, \ldots, (n-1)/n, \ldots\} \cup \{0, 1\}$, with $p < q$. The rationale for Definition 18 and the fractional units it employs is given by the next theorem.

THEOREM 10. $\mathbf{POPPERFN}[p, q] \subseteq \mathbf{Q}[p, q]$, *for p and q as specified above.*

Proof. The proof of this theorem explains the origin of the rules in Definition 18. The proof depends only on obvious characteristics of probability functions. First, the probabilistic soundness of the rules of $\mathbf{Q}$ guarantees that $\mathbf{POPPERFN}[0, 1] \subseteq \mathbf{Q}[0, 1]$. Also, we already know that $\mathbf{POPPERFN}[1, 1] = \mathbf{POPPERFN}[1] = \mathbf{R} = \mathbf{Q}[1, 1]$. Regarding $\mathbf{Q}[1/(n+1), 1]$, consider the following theorem for Popper Functions: if $C \models \neg(B_i \& B_j)$ (for $i \neq j$), then $P[B_1 \mid C] + \cdots + P[B_{n+1} \mid C] \leqslant 1$ or $P[\neg C \mid C] = 1$. So not all $P[B_i \mid C] \geqslant r > 1/(n+1)$ unless $P[\neg C \mid C] = 1$. Thus, any member of $\mathbf{POPPERFN}[1/(n+1), 1]$ must satisfy the $\mathbf{Q}[1/(n+1), 1]$ rule. Similar observations connect levels of support for Popper Functions to the other rules in Definition 18. $\mathbf{Q}[(n-1)/n, 1]$: if $C \models (B_1 \vee B_2 \vee \cdots \vee B_n)$ and $P[\neg B_i \mid C] \geqslant r > (n-1)/n$, then $P[B_i \mid C] < 1/n$ (unless $P[\neg C \mid C] = 1$), so $1 = P[B_1 \vee B_2 \vee \cdots \vee B_n \mid C] \leqslant P[B_1 \mid C] + \cdots + P[B_n \mid C] < 1$ unless $P[\neg C \mid C] = 1$. $\mathbf{Q}[0, n/(n+1)]$: if $C \models \neg(B_i \& B_j)$ for $B_1, \ldots, B_{n+1}$, and if $P[\neg B_i \mid C] < r \leqslant n/(n+1)$, then $1 \geqslant P[B_1 \mid C] + \cdots + P[B_{n+1} \mid C] > (n+1) \times (1/(n+1)) = 1$. $\mathbf{Q}[0, 1/n]$: if $C \models (B_1 \vee B_2 \vee \cdots \vee B_n)$, and $P[B_i \mid C] < r \leqslant 1/n$, then $1 = P[B_1 \vee B_2 \vee \cdots \vee B_n \mid C] \leqslant P[B_1 \mid C] + \cdots + P[B_n \mid C] < n/n = 1$. □

Theorem 10 shows that the rules of Definition 18 are sound for the conditionals defined by $\mathbf{POPPERFN}[p, q]$. However, there are conditionals in $\mathbf{Q}[p, q]$ that cannot be extended into members of $\mathbf{POPPERFN}[p, q]$. The way in which rules for conditional probabilities assign numbers to sentences imposes important *ordering relations* on sentences that some conditionals in $\mathbf{Q}[p, q]$ do not heed. In the next subsection I will specify rules that, when satisfied by a conditional in $\mathbf{Q}$, suffice to guarantee that it exhibits the salient ordering relations. In the two subsections after the next I will show how to extend these primitive orderings to characterize just those conditionals $\mathbf{Q}[p, q]$ that belong to $\mathbf{POPPERFN}[p, q]$.

4.2. *Orderings Imposed by Some Conditionals in* $\mathbf{Q}$

Think of the sentences of language L as representing possibilities, sets of possible states of affairs or possible worlds. Intuitively some possibilities are more weighty or more likely than others. Only those conditionals in $\mathbf{Q}$ that systematically reflect the degree to which some possibilities are (treated by the conditional as) more weighty than others will behave precisely like probabilities. It will prove useful to explicate the weightiness attributed to sentences by a conditional in terms of sets of mutually exclusive sentences called *partitions*.

DEFINITION 19. A finite set U of two or more consistent, mutually exclusive sentences (i.e. consistent sentences that are pairwise inconsistent) in S_L will be called a *partial partition* of S_L. The members of U are the partial partition's *elements*. If U is also exhaustive (i.e. if the disjunction of elements of U is logically true), then U is a (*complete*) *partition* of S_L. Sentence B of S_L will be called *representable in partial partition* U just in case either $\models B$ or $\models \neg B$ or, for some subset $\{C_1, \ldots, C_n\}$ of U, $\models B \equiv (C_1 \vee \cdots \vee C_n)$. If B is representable by the set of n elements $\{C_1, \ldots, C_n\}$ of U, define $\#_U[B] = n$; if $\models \neg B$, define $\#_U[B] = 0$; if $\models B$, define $\#_U[B] =$ the number of elements of U.

The rules in the next definition specify a "weightiness ordering" associated with conditionals in $\mathbf{Q}$. This "ordering" is not necessarily transitive for conditionals in $\mathbf{Q}$, but should be so for probabilities.

DEFINITION 20. For each conditional $\rightarrow$ in $\mathbf{Q}$ and each partial partition U of S_L, define a relation $\geqslant_U$ on sentences representable in U as follows:

$A \geqslant_U B$ iff for all C, D representable in U such that $C \models D$ and $\models \neg(A\&D)\&\neg(B\&D)$, the following 3 rules hold:

$$(1)\ \frac{D \vee A \rightarrow C}{D \vee B \rightarrow C} \qquad (2)\ \frac{D \vee B \rightarrow C \vee B}{D \vee A \rightarrow C \vee A}$$

$$(3)\ \frac{D \vee B \vee A \rightarrow C \vee B}{D \vee B \vee A \rightarrow C \vee A}.$$

The relations $>_U$ and $\sim_U$ for a given $\rightarrow$ are defined as follows:

$A >_U B$ iff $A \geqslant_U B$ and not $B \geqslant_U A$;

$A \sim_U B$ iff $A \geqslant_U B$ and $B \geqslant_U A$.

Roughly '$A \geqslant_U B$' says that "*A is at least as weighty as B*" for the conditional in **Q** associated with $\geqslant_U$. Notice that if all conditional assertions in rules 1–3 were replaced with conditional probabilities at a level of support r, then these rules would require that A is at least as probable as B.

In general there is no reason to expect a relation $\geqslant_U$ associated with an arbitrary conditional in **Q** to be either transitive or connected. But if a conditional in **Q** is to be truly probability-like, then an *at-least-as-weighty-as* relation, $\geqslant_U$, associated with it will have to be a weak order (i.e. a transitive and connected ordering). For, every probability function imposes this kind of ordering on sentences in conformity with their probabilistic weights, an ordering that satisfies the rules for a *qualitative probability relation* (for treatments of qualitative probability see (Savage 1954), (Krantz, Luce, Suppes, Tversky 1971), (Suppes, Krantz, Luce, Tversky 1989), and (Narens 1985)). I won't delve into the theory of qualitative probability here. Rather, I will take a more direct approach to the specification of conditionals in $\mathbf{Q}[p,q]$ that the belong to $\mathbf{POPPERFN}[p,q]$.

A conditional in **Q** behaves like conditional probability on those parts of its language that are representable by a partial partition that is *uniform enough*, a partial partition in which all pairs of elements are $\sim_U$ equivalent. This idea is made precise by the next definition and the following two theorems.

DEFINITION 21. A partial partition U is *uniform enough* for $\rightarrow$ iff for each pair of its *elements* A, B, $A \sim_U B$.

Notice that for a partial partition to be uniform enough its elements need only satisfy the rules in Definition 20 relative to sentences representable in the partition. The elements of a uniform enough partial partition need only behave as though they are "equally weighty" relative to the sentences representable in the partition. They need not be "equally weighty" relative to a broader class of sentences or a finer partition of the language. On those parts of the language where there is a uniform enough partial partition for a conditional in **Q**, the conditional behaves like a conditional probability function at a support level.

THEOREM 11. *Let $\rightarrow$ be in* **Q**, *and let U be a uniform enough partial partition for $\rightarrow$. For all sentences A, B, C, D in S_L representable in U, if $B \rightarrow A$ and $D \not\rightarrow C$, then $\#_U[A\&B] \div \#_U[B] > \#_U[C\&D] \div \#_U[D]$.*

Proof. First notice that for any conditional, $\rightarrow$, in **O** that has a partial partition U, the number of elements of U that occur as disjuncts in (the representation in U of) the antecedent of a conditional assertion, and the number of those elements that occur in (the representation in U of) the consequent completely determines whether the conditional assertion holds. That is, from the definition of $\geqslant_U$ together with the fact that each element of U bears $\sim_U$ to every other it follows from rules of **O** that for *some* distinct elements $\{E_1, \ldots, E_n\}$ of U and $m \leqslant n$, $E_1 \vee \cdots \vee E_n \rightarrow E_1 \vee \cdots \vee E_m$ *if and only if* for *every* set of n distinct elements $\{G_1, \ldots, G_n\}$ of U, $G_1 \vee \cdots \vee G_n \rightarrow G_1 \vee \cdots \vee G_m$. So, whether or not a conditional holds depends only on the number of partition elements in the antecedent and consequent. Thus, when $E_1 \vee \cdots \vee E_n \rightarrow E_1 \vee \cdots \vee E_m$ holds we may abbreviate this by the expression '$n \rightarrow m$', and we may write '$n \not\rightarrow m$' to abbreviate that $E_1 \vee \cdots \vee E_n \not\rightarrow E_1 \vee \cdots \vee E_m$.

Let A, B, C, D, and U satisfy the antecedent of the theorem. Notice, $B \rightarrow A$ iff $B \rightarrow B\&A$, and $D \not\rightarrow C$ iff $D \not\rightarrow D\&C$. Let B be represented in U by $B_1 \vee \cdots \vee B_b$ and $B\&A$ be represented in U by $B_1 \vee \cdots \vee B_a$ for some $a \leqslant b$. Similarly, let D be represented by $D_1 \vee \cdots \vee D_d$, and let $D\&C$ be represented by $D_1 \vee \cdots \vee D_c$ for some $c < d$. So, $B_1 \vee \cdots \vee B_b \rightarrow B_1 \vee \cdots \vee B_a$ and $D_1 \vee \cdots \vee D_d \not\rightarrow D_1 \vee \cdots \vee D_c$. We want to show that $a \div b > c \div d$, so *assume for reductio* that $c \div d \geqslant a \div b$. Thus, we have $b \rightarrow a$, $d \not\rightarrow c$, and we are assuming that $b \div a \geqslant d \div c > 1$, and we're looking for a contradiction.

Now, the strategy of the proof is to apply an iterative process that strips elements of U from $b \rightarrow a$ and $d \not\rightarrow c$, leaving ever shorter disjunctions of elements of U on either side of '$\rightarrow$' and '$\not\rightarrow$'. To specify this iterative process there are just three kinds of cases to consider:

(1) Suppose $d > b$. Then $c > a$, so (Weak Or) implies that $d - b \not\rightarrow c - a$, with $b \div a \geqslant (d - b) \div (c - a) > 1$.

(2) Suppose $b \geqslant d$. Then either $a > c$ or $c \geqslant a$.
(2.1) If $a > c$, then $b > d$, so (Negation Rationality) implies $b - d \rightarrow a - c$, with $(b - d) \div (a - c) \geqslant d \div c > 1$.
(2.2) If $c \geqslant a$, then $b \rightarrow c$ (and $d \not\rightarrow c$, so $b > d$), so $b - d \rightarrow c - c$ (where $c - c$ is $\mathbf{F}$).

Case 2.2 is impossible since members of U cannot bear $\rightarrow$ to $\mathbf{F}$ (this follows from '$D \not\rightarrow C$'). Now notice that cases (1) and (2.1) can be iterated: if case (1) applied and yielded '$d - b \not\rightarrow c - a$', then using '$d - b \not\rightarrow c - a$' in place of '$d \not\rightarrow c$' again apply either case (1) or case (2) (depending on whether $d - b \geqslant b$ or $b > d - b$); if case (2.1) applied and yielded '$b - d \rightarrow a - c$', then using '$b - d \rightarrow a - c$' in place of '$b \rightarrow a$' again apply either case (1) or case (2) (depending on whether $b - d \geqslant d$ or $d > b - d$). The iteration of this process strips elements of U from $b \rightarrow a$ and $d \not\rightarrow c$, leaving ever shorter disjunctions of elements of U on either side of '$\rightarrow$' and '$\not\rightarrow$' until an instance of case (2.2) arises to produce a contradiction. □

Consider any conditional, $\rightarrow$, in $\mathbf{Q}$ relative to which a partial partition U is uniform enough. Define the function P_U as follows: for A, B representable in U, $P_U[A \mid B] = \#_U[A\&B] \div \#_U[B]$ or $\models B \supset \mathbf{F}$ and $P_U[A \mid B] = 1$. Clearly P_U is a classical conditional probability function *on the algebra of sentences representable in* U. The previous theorem implies that there must be some real number r such that for all sentences A and B representable in U, $P_U[A \mid B] \geqslant r$ if and only if $B \rightarrow A$. The next theorem establishes that if $\rightarrow$ also satisfies the rules for $\mathbf{Q}[p, q]$, then the value of a real number r that divides support from non-support must lie between p and q.

THEOREM 12. *Let $\rightarrow$ be a member of $\mathbf{Q}[p, q]$, and let U be a uniform enough partial partition for $\rightarrow$. Then for some r such that $p < r \leqslant q$, for all A, B representable in U, $\#_U[A\&B] \div \#_U[B] \geqslant r$ iff $B \rightarrow A$.*

Proof. $\mathbf{Q}[p, q] \subseteq \mathbf{Q}$ and Theorem 11 implies that for some r, $0 < r \leqslant 1$, $\#_U[A\&B] \div \#_U[B] \geqslant r$ iff $B \rightarrow A$. So we just need to show that $p < r \leqslant q$. The rules of $\mathbf{O}$ permit us to restrict attention to sentences A and B such that $A \models B$, and to further restrict attention to their respective representations in U of form $G_1 \vee \cdots \vee G_b$ and $G_1 \vee \cdots \vee G_a$, for $b \geqslant a$. Thus, it will suffice to prove the following cases:

(1) if $G_1 \vee \cdots \vee G_b \rightarrow G_1 \vee \cdots \vee G_a$, then $a/b > p$ (we may define r as the smallest such ratio $a/b > p$);
(2) if $G_1 \vee \cdots \vee G_b \not\rightarrow G_1 \vee \cdots \vee G_a$, then $a/b < q < 1$.
(1) Suppose that $G_1 \vee \cdots \vee G_b \rightarrow G_1 \vee \cdots \vee G_a$ and $p \geqslant a/b$.

(1.1) Let $p = 1/(n+1)$ for some integer $n \geqslant 2$.

Then $1/(n+1) \geqslant a/b$, so $b \geqslant (n+1)a \geqslant n+1$. Thus, $G_1 \vee \cdots \vee G_b$ can be subdivided into $n+1$ segments with at least a disjuncts from U in each segment. From the $\sim_U$ equivalence of elements of U, $G_1 \vee \cdots \vee G_b$ must bear $\rightarrow$ to each of these segments. Then by the rule for $\mathbf{Q}[1/n+1, 1]$ and the logical incompatibility of each pair of segments we have $G_1 \vee \cdots \vee G_b \rightarrow \mathbf{F}$. But then $G_i \rightarrow \mathbf{F}$ for all of the G_i in U (by repeated application of (Negation Rationality)), which contradicts the definition of U.

(1.2) Let $p = (n-1)/n$ for some integer $n \geqslant 2$.

Then $(n-1)/n \geqslant a/b$, so $n(b-a) \geqslant b > a$. Thus, $G_1 \vee \cdots \vee G_b$ can be subdivided into n or fewer segments, $D_1, \ldots, D_k$, with $b-a$ or fewer elements of U in each segment. Let D_i be any one of these segments. Then $G_1 \vee \cdots \vee G_b \rightarrow \neg D_i$ (since $G_1 \vee \cdots \vee G_b \rightarrow G_1 \vee \cdots \vee G_a$, so $G_1 \vee \cdots \vee G_b \rightarrow (G_1 \vee \cdots \vee G_a) \& \neg(G_{a+1} \vee \cdots \vee G_b)$, so $G_1 \vee \cdots \vee G_b \rightarrow \neg(G_{a+1} \vee \cdots \vee G_b)$, and similarly for any segment with $b-a$ or fewer elements of U). $G_1 \vee \cdots \vee G_b \models D_1 \vee \cdots \vee D_k$. So, $G_1 \vee \cdots \vee G_b$ logically entails this disjunction of strings supplemented with enough repetitions of the first string to make exactly n disjoined strings. Then, by the rule for $\mathbf{Q}[(n-1)/n, 1]$, $G_1 \vee \cdots \vee G_b \rightarrow \mathbf{F}$, which leads to $G_i \rightarrow \mathbf{F}$ for each G_i, which contradicts the definition of U.

(2) Suppose $G_1 \vee \cdots \vee G_b \not\rightarrow G_1 \vee \cdots \vee G_a$ and $1 > a/b \geqslant q$.

(2.1) Let $q = n/(n+1)$ for some integer $n \geqslant 2$.

The $a/b \geqslant n/(n+1)$, so $b \geqslant (n+1) \times (b-a) \geqslant (n+1)$. Thus, $G_1 \vee \cdots \vee G_b$ can be subdivided into $n+1$ mutually inconsistent segments with at least $b-a$ disjuncts from U in each. The rule for $\mathbf{Q}[0, n/(n+1)]$ yields $G_1 \vee \cdots \vee G_b \rightarrow \neg D$ for at least one of these $(n+1)$ segments, D. It follows that $G_1 \vee \cdots \vee G_b \rightarrow (G_1 \vee \cdots \vee G_b) \& \neg D$; and $(G_1 \vee \cdots \vee G_b) \& \neg D$ is logically equivalent to some disjunction $G_i \vee \cdots \vee G_k$ that contains a or fewer distinct disjuncts from U. Then, from the $\sim_U$ equivalence of the G_i, it follows that $G_1 \vee \cdots \vee G_b \rightarrow G_1 \vee \cdots \vee G_a$, a contradiction.

(2.2) Let $q = 1/n$ for some integer $n \geqslant 2$.

Then, $a/b \geqslant 1/n$, so $na \geqslant b > a$. Thus, $G_1 \vee \cdots \vee G_b$ can be subdivided into n or fewer segments, $D_1, \ldots, D_k$, with a or fewer elements of U in each segment. Then $G_1 \vee \cdots \vee G_b \not\rightarrow D_i$ for each D_i. This contradicts the rule for $\mathbf{Q}[0, 1/n]$. □

As a consequence of Theorem 12 it is easy to illustrate how conditionals in $\mathbf{Q}[p, q]$ deal with the *lottery paradox* for various levels of p and q.

Consider a fair lottery in which 100 tickets are sold. Let C describe relevant background conditions about the nature of the lottery, and let each sentence W_i, for $1 \leqslant i \leqslant 100$, say "ticket i will win". The condition that exactly one ticket will win is captured by the formal requirements that $C \models W_i \vee \cdots \vee W_{100}$ and $C \models \neg(W_i \& W_j)$ for distinct tickets i and j. Let $U = \{C\&W_1, \ldots, C\&W_{100}\}$ be a uniform enough partial partition on the conditional $\rightarrow$. A fair lottery should be represented by a conditional for which the "weightiness" of the claim that a given ticket will win is the same for each ticket. The *uniformity* of the partial partition U for $\rightarrow$ formalizes this condition.

Theorem 12 implies that for $\rightarrow$ in $\mathbf{Q}[99/100, 100/101]$, $C \not\rightarrow \neg W_i$ for each of the tickets, since $\#_U[W_1\&C \vee \cdots \vee W_{99}\&C] \div \#_U[C] = 99/100$. But if $\rightarrow$ is a member of $\mathbf{Q}$ that corresponds to a less stringent support level, say $\mathbf{Q}[98/99, 99/100]$, then $C \rightarrow \neg W_i$ holds for each ticket. At this level, though, $C \not\rightarrow \neg W_i\&\neg W_j$ holds for each distinct pair of tickets i and j. In general it can be shown that if $\rightarrow$ is in $\mathbf{Q}[(n-1)/n, n/(n+1)]$ (alternatively, $\mathbf{Q}[1/(n+1), 1/n]$) for any integer n from 99 through 2, then there is an integer k, where $(100 \div (n+1)) - 1 < k < 100 \div n$ (alternatively, $100 \times ((n-1) \div n) - 1 < k < 100 \times (n \div (n+1))$), such that $C \rightarrow \neg W_i \& \cdots \& \neg W_j$ holds for all conjunctions of k or fewer tickets but $C \not\rightarrow \neg W_i \& \cdots \& \neg W_j$ holds for all conjunctions of more than k tickets.

Uniform enough partial partitions will provide a qualitative way to precisely identity the conditionals in $\mathbf{Q}[p, q]$ that belong to $\mathbf{POPPERFN}[p, q]$.
The next two subsections will show how.

4.3. *Probabilistic Representations of Conditionals on Finite Languages*

I will restrict attention now to languages containing any finite number of sentence letters. I am not yet sure how to extend the following results to infinite languages, but I suspect some adaptation of the work on qualitative probability on the non-standard reals will do the trick. For any finite language L let $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$ and $\mathbf{Q}_{\mathrm{L}}[p, q]$ to be the sets of conditionals on L that satisfy the definitions of $\mathbf{POPPERFN}[p, q]$ and $\mathbf{Q}[p, q]$, respectively.

Theorem 10 established that all conditionals in $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$ satisfy the rules of $\mathbf{Q}_{\mathrm{L}}[p, q]$. How may we characterize those members of $\mathbf{Q}_{\mathrm{L}}[p, q]$ that impose sufficient ordering relations on sentences to qualify for membership in $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$? In this subsection I will answer this question for an important part of $\mathbf{POPPERFN}_{\mathrm{L}}[p,q]$; the next subsection extends the answer to all of $\mathbf{POPPERFN}_{\mathrm{L}}[p,q]$.

The conditional probability functions that derive from classical probability theory were specified in Definition 2. Let's call the set of these classical conditional probability functions **CLASSCOND**. It is easily verified that $\mathbf{CLASSCOND} \subset \mathbf{POPPERFN}$; the members of **CLASSCOND** are just those Popper Functions that contain only one finite rank, and also rank ∞. Thus, the set of conditionals

$$\mathbf{CLASSCOND}_{\mathrm{L}}[p,q],$$

when defined in the obvious way, will be a subset of $\mathbf{POPPERFN}_{\mathrm{L}}[p,q]$.

DEFINITION 22. $\rightarrow\ \in \mathbf{CLASSCOND}[p,q]$ iff there is a $P \in \mathbf{CLASSPROB}$ and some real number r,

$$p < r \leqslant q, \quad \text{such that} \rightarrow\ = \{\langle B, A\rangle \mid P[A \mid B] \geqslant r\}.$$

Theorem 10 implies that $\mathbf{CLASSCOND}[p,q] \subseteq \mathbf{Q}[p,q]$.

Clearly, if $\rightarrow$ is a conditional in $\mathbf{Q}_{\mathrm{L}}[p,q]$ (for finite L) and the set of state descriptions of L is a uniform enough *complete* partition of L for $\rightarrow$, then $\mathbf{CLASSCOND}_{\mathrm{L}}[p,q]$ will contain $\rightarrow$; this follows directly from Theorem 12. But for most conditionals in $\mathbf{Q}_{\mathrm{L}}[p,q]$ the state descriptions will not constitute a *uniform enough* partition. However, many of these conditionals satisfy the rules of Definition 20 in a coherent enough way to be in $\mathbf{CLASSCOND}_{\mathrm{L}}[p,q]$. The following definition picks out just the right class.

DEFINITION 23. For any finite language L, $\rightarrow\ \in \mathbf{Q}^*_{\mathrm{L}}[p,q]$ iff there is a finite extension of L, L^+, and an extension of $\rightarrow$ to L^+, $\rightarrow^+$ (i.e. $\rightarrow^+$ agrees with $\rightarrow$ on sentences of L), such that:

(1) $\rightarrow^+ \in \mathbf{Q}_{\mathrm{L}^+}[p,q]$;
(2) there is a uniform enough partial partition U for $\rightarrow^+$ such that each state description in L is logically equivalent to a disjunction composed only of elements from U and sentences from $F = \{B \mid B \rightarrow^+ \mathbf{F}\}$.

The next theorem shows that the conditionals in $\mathbf{Q}^*_{\mathrm{L}}[p,q]$ are precisely those that are representable by conditional probability functions. These conditional probability functions employ a probabilistic level of support r, for $p < r \leqslant q$, that reflects the notion of support engendered in the $\mathbf{Q}_{\mathrm{L}}[p,q]$ rules.

THEOREM 13. *For any finite* L, $\mathbf{Q}^*_{\mathrm{L}}[p,q] = \mathbf{CLASSCOND}_{\mathrm{L}}[p,q]$.

Proof. (1) Suppose $\rightarrow\ \in \mathbf{Q}^*_{\mathrm{L}}[p,q]$, and let $\rightarrow^+$ be the extension of $\rightarrow$ to L^+ described in Definition 23. Some sentence C in L is logically equivalent to a disjunction of elements from both U and F, or even from F alone. So, extend the definition of $\#_U$ as follows: $\#_U[C] =$ the number of elements from U in a disjunction of elements from U and F that is logically equivalent to C. For all A, B in L define $P[A \mid B] = \#_U[A\&B] \div \#_U[B]$, or 1 if $B \rightarrow \mathbf{F}$. P is a classical conditional probability function on language L. And, by Theorem 12, there is an r, $p < r \leqslant q$, such that $P[A \mid B] \geqslant r$ just in case $B \rightarrow A$.

(2) Conversely, suppose $\rightarrow\ \in \mathbf{CLASSCOND}_{\mathrm{L}}[p,q]$, and let P and r be the classical conditional probability function and support level (for $p < r \leqslant q$) that define $\rightarrow$. Since the language L is finite there exists a classical conditional probability function P' (with only rational number probability values) that behaves just like P in the sense that for all sentences A, B, $P'[A \mid B] \geqslant r$ <u>iff</u> $P[A \mid B] \geqslant r$ (and for which the unconditional probabilities are also rational numbers). Let $\{D_1, \ldots, D_n\}$ be the state descriptions of L. P' assigns each state description either 0 or a positive rational number. Let s be the largest positive rational number such that for every D_i, $P'[D_i]$ is an integer multiple of s. The language L can now be extended to a language L^+ with enough new sentence letters to subdivide all state descriptions D_i into new state descriptions, where each new state description is either assigned probability s or probability 0 by a function P^+ that agrees with P' on L (since the state description probabilities for L^+ can clearly be assigned so as to sum to the values of $P'[D_i]$ for each D_i). The new state descriptions that have positive probabilities provide a uniform enough partial partition U for the conditional $\rightarrow^+$ (where $B \rightarrow^+ A$ <u>iff</u> $P^+[A \mid B] \geqslant r$). And $\rightarrow^+$ satisfies the rules of $\mathbf{Q}_{\mathrm{L}}[p,q]$. □

The requirements for membership in $\mathbf{Q}^*_{\mathrm{L}}[p,q]$ are rather modest constraints on conditionals that belong to $\mathbf{Q}_{\mathrm{L}}[p,q]$ already. Let $\{D_1, \ldots, D_n\}$ be the set of state descriptions of L. If they are all $\sim$ equivalent for conditional $\rightarrow$ (except those in F, i.e. those that bear $\rightarrow$ to $\mathbf{F}$), then membership in $\mathbf{Q}^*_{\mathrm{L}}[p,q]$ is assured. If the state descriptions are not $\sim$ equivalent for $\rightarrow$, extend L to language L^+ containing an exclusive and exhaustive set of "bets" $\{W_1, \ldots, W_m\}$, each compatible with the state description of L (except those in F). The "bets" W_j in the extended language might, for example, each claim that a different ticket (or no ticket) will win some lottery that has absolutely nothing to do with the subject matter of language L. For each state description D_i of L (except those in F), let U contain mutually exclusive sentences of form $D_i\&(W_j \vee \cdots \vee W_k)$.

Effectively, the weight of each state description is diluted by conjoining it to some disjunction of conjectures about the lottery outcome. Each state description of L is subdivided into less weighty sentences. If for some extension $\rightarrow^+$ of $\rightarrow$ there is such a partial partition U that subdivides the state descriptions of L into $\sim_U$ equivalent parts, then $\rightarrow$ is in $\mathbf{Q}^*_{\mathrm{L}}[p, q]$.

4.4. *Conditionals from* Q *that Contain Conditionals in* R

The set of conditionals $\mathbf{Q}^*_{\mathrm{L}}[p, q] = \mathbf{CLASSCOND}_{\mathrm{L}}[p, q]$ is a proper subset of $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$. We will need to expand the definition of $\mathbf{Q}^*_{\mathrm{L}}[p, q]$ if we are to capture $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$ precisely. $\mathbf{Q}^*_{\mathrm{L}}[p, q]$ comes up short because a partial partition for $\rightarrow$ is not *uniform enough* unless for *every pair* of sentence A and B, A can be subdivided into parts, $A\&C_i$, each of which is no weightier (in terms of $\geqslant$ for $\rightarrow$) than B (except when $B \rightarrow \mathbf{F}$). But conditionals in $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$ may arise from Popper Functions that contain two or more finite ranks, and no two finite ranks can be represented by a single uniform enough partial partition.

The obvious way to extend the definition of $\mathbf{Q}^*_{\mathrm{L}}[p, q]$ to capture all of $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$ is to consider the set of conditionals in $\mathbf{Q}$ that are ranked by the $\gg$ relation of Definition 13. I'll call this set $\mathbf{QR}$. Then define $\mathbf{Q}^*_{\mathrm{L}}[p, q]$ to be the set of conditionals in $\mathbf{QR}$ that can be extended to possess uniform enough partial partitions *at each rank*. These will turn out to be just the conditionals in $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$. The remainder of this section executes the details of this idea.

Definition 13 specified ordering relations, $\gg$, that conditionals in $\mathbf{O}$ impose on sentences. Theorem 6 showed these to be strict partial orders (i.e. transitive and asymmetric). The associated symmetric relations $\approx$ are reflexive but may not in general be transitive. However, when $\approx$ is transitive the associated symmetric relation $\underline{\gg}$ is automatically a weak order. So let's restrict attention to conditionals for which $\approx$ is transitive. For these conditionals the relations $\underline{\gg}$ will rank the sentences of L.

DEFINITION 24. $\rightarrow\, \in \mathbf{QR}[p, q]$ (more generally, $\rightarrow\, \in \mathbf{OR}$, and $\rightarrow\, \in \mathbf{QR}$) iff $\rightarrow\, \in \mathbf{Q}[p, q]$ ($\rightarrow\, \in \mathbf{O}$, $\rightarrow\, \in \mathbf{Q}$) and satisfies the following condition: if $A \approx B$ and $B \approx C$, then $A \approx C$. For $\rightarrow\, \in \mathbf{QR}[p, q]$ (or $\mathbf{OR}$ or $\mathbf{QR}$) define the rank function and an associated conditional $-\!\!>$ in $\mathbf{R}$ as follows:

$$\begin{aligned}
\mathrm{rank}[A] &= 1 \text{ \underline{iff} } \mathbf{T} \approx A; \\
\mathrm{rank}[A] &= i + 1 \text{ \underline{iff} } A \not\rightarrow \mathbf{F},\ \mathrm{rank}[A] \underline{\not<} i, \text{ and for all } B, \\
&\qquad \mathrm{rank}[B] \leqslant i \text{ or } B \rightarrow \mathbf{F} \text{ or } A \underline{\gg} B;
\end{aligned}$$

rank$[A] = \infty$ iff $A \rightarrow F$.

$B{-}{>}A$ iff for all $C \in \mathrm{SD}\{A, B\}$ such that rank$[C] =$ rank$[B]$, if $C \models B$ then $C \models A$.

Theorem 9 and the discussion following it established that the ranking specified by a weak order $\geqslant$ on state descriptions completely characterizes a conditional in **R**. Definition 24 takes advantage of this fact to uncover a conditional $->$ in **R** superimposed on each conditional $\rightarrow$ in **QR**.

Now let us return to the question of primary interest: for a finite language L, how may we characterize those members of $\mathbf{QR}_{\mathrm{L}}[p, q]$ that impose sufficient ordering relations on sentences to qualify for membership in $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$? The next definition and theorem extend the answer given in the previous subsection (by Definition 23 and Theorem 13).

DEFINITION 25. For a finite language L, $\rightarrow \in \mathbf{QR}^*_{\mathrm{L}}[p, q]$ iff there is a finite extension of L, L^+, and an extension of $\rightarrow$ to L^+, $\rightarrow^+$, with the following properties:

(1) $\rightarrow^+ \in \mathbf{QR}_{\mathrm{L}^+}[p, q]$;
(2) for each finite rank i of $\rightarrow^+$, there is a uniform enough partial partition, U_i, containing only elements of rank i;
(3) for every state description D in L, either the rank of D is ∞, or there is a finite rank i such that D is logically equivalent to a disjunction of elements from U_i and sentences of rank ∞.

The conditionals in $\mathbf{QR}^*_{\mathrm{L}}[p, q]$ turn out to be precisely those that correspond to the probability greater than r_i parts (for each rank i) of Popper Functions, as spelled out in Definition 16.

THEOREM 14. *For any finite* L, $\mathbf{QR}^*_{\mathrm{L}}[p, q] = \mathbf{POPPERFN}_{\mathrm{L}}[p, q]$.

Proof. Observe that for conditionals in either $\mathbf{QR}^*_{\mathrm{L}}[p, q]$ or $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$, a conditional assertion $B \rightarrow A$ holds just in case another conditional $D \rightarrow C$ holds for D the disjunction of the highest ranking state descriptions that logically entail B and C the disjunction of these state descriptions that also entail $B\&A$. The present theorem then follows from the application of the proof of Theorem 13 to each rank of the conditionals in $\mathbf{QR}^*_{\mathrm{L}}[p, q]$ and in $\mathbf{POPPERFN}_{\mathrm{L}}[p, q]$. □

Theorem 14 says that the logic of probabilistic support at levels between p and q is *soundly* and *completely* captured by the rules for conditionals in $\mathbf{QR}^*_{\mathrm{L}}[p,q]$. One important implication is that logics for nonmonotonic conditionals and conditional probabilities may be coherently integrated in computer based defeasible reasoning systems. More generally, Theorem 14 suggests that qualitative and probabilistic approaches to uncertain knowledge share a common logic.

5. CONCLUSION

Most of the inferences we commonly make are defeasible, and nonmonotonic conditionals provide a way to model significant features of the logic of defeasible reasoning. The nonmonotonic logic embodied by conditional probabilities has long been our best model of defeasible support for beliefs by evidence, a model in which *support* for a belief reflects the *probability* that it is true. Each nonmonotonic conditional in a family $\mathbf{QR}^*[p,q]$ behaves precisely like a conditional probability function at some level of probabilistic support between p and q. These conditionals provide a logically sound qualitative model of the support and revision of beliefs due to evidence, a model in which *support* is an indicator of the probability that the belief is true. Conversely, any nonmonotonic conditionals that violates rules of $\mathbf{QR}^*[p,q]$ must conflict with the logic of conditional probability. Thus, any conditional that does not belong to some $\mathbf{QR}^*[p,q]$ cannot effectively represent the *probable truth* (at some level) of "supported" beliefs.

NOTES

[1] Chris Swoyer provided very helpful comments on drafts of this paper.

[2] See the investigations of Adams (1966, 1975), Stalnaker (1970), Harper (1975), Pollock (1976), Pearl (1988), and Lehmann and Magidor (1992).

[3] Both indicative and subjunctive (or counterfactual) conditionals are nonmonotonic. Among the most influential treatments of these conditionals are those cited in the previous note, and the following: (Stalnaker 1968), (Lewis 1973), (Nute 1980), (Kraus, Lehmann, Magidor 1990); and also the papers collected in (Harper, Stalnaker, Pearce 1981), (Reinfrank, De Kleer, Ginsberg, Sandewall 1989), and (Jackson 1991). Nute (1984) provides an excellent overview of conditional logics.

[4] Harper (1975) explored the connection between Popper Functions and personalist probability, and first recognized a connection between Popper Functions and conditionals. Field (1977) developed Popper Functions into a semantics for first-order logic. See Leblanc (1979, 1983) and van Fraassen (1981) for further developments of probabilistic semantics for first-order logic.

REFERENCES

Adams, E., 1966: 'Probability and the Logic of Conditionals', J. Hintikka and P. Suppes (eds), *Aspects of Inductive Logic* (North-Holland, Amsterdam), 265–316.

Adams, E., 1975: *The Logic of Conditionals* (Reidel, Dordrecht).

Field, H., 1977: 'Logic, Meaning, and Conceptual Role', *Journal of Philosophy* **74**, 379–409.

Harper, W., 1975: 'Rational Belief Change, Popper Functions and Counterfactuals', *Synthese* **30**, 221–262.

Harper, W., R. Stalnaker, and G. Pearce (eds), 1981: *Ifs: Conditionals, Belief, Decision, Chance and Time* (Reidel, Dordrecht).

Jackson, F. (ed.), 1991: *Conditionals* (Oxford Univ. Press, Oxford).

Krantz, D., R. Luce, P. Suppes, and A. Tversky, 1971: *Foundations of Measurements*, Vol. I (Academic Press, New York).

Kraus, S., D. Lehmann, and M. Magidor, 1990: 'Nonmonotonic Reasoning, Preferential Models and Cumulative Logics', *Artificial Intelligence* **44**(1–2), 167–207.

Leblanc, H., 1979: 'Probabilistic Semantics for First-order Logic', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **25**, 497–509.

Leblanc, H., 1983: 'Alternatives to Standard First-order Semantics', in D. Gabbay and F. Guenthner (eds), *Handbook of Philosophical Logic*, Vol. I (Reidel, Dordrecht), Chapter I.3, 189–247.

Lehmann, D. and M. Magidor, 1992: 'What Does a Conditional Knowledge Base Entail?' *Artificial Intelligence* **55**, 1–60.

Lewis, D., 1973: *Counterfactuals* (Harvard Univ. Press, Cambridge, MA).

Narens, L., 1985: *Abstract Measure Theory* (MIT Press, Cambridge, MA).

Nute, D., 1980: *Topics in Conditional Logic* (Reigel, Dordrecht).

Nute, D., 1984: 'Conditional Logic', in D. Gabbay and F. Guenthner (eds), *Handbook of Philosophical Logic*, Vol. II (Reidel, Dordrecht), Chapter II.8, 387–439.

Pearl, J., 1988: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Reasoning* (Morgan Kaufmann, San Mateo, CA).

Pollock, J., 1976: *Subjunctive Reasoning* (Reidel, Dordrecht).

Popper, K., 1968: *The Logic of Scientific Discovery* (Harper & Row, New York).

Reinfrank, M., J. De Kleer, M. L. Ginsberg, and E. Sandewall (eds), 1989: *Non-Monotonic Reasoning, Lecture Notes in Artificial Intelligence* (Springer-Verlag, Berlin).

Savage, L., 1954: *The Foundations of Statistics* (Wiley, New York; Dover, 1972).

Stalnaker, R., 1968: 'A Theory of Conditionals', N. Rescher (ed.), *Studies in Logical Theory, American Philosophical Quaterly*, Monograph Series 2 (Blackwell, Oxford), 98–112.

Stalnaker, R., 1970: 'Probability and Conditionals', *Philosophy of Science* **37**, 64–80.

Suppes, P., D. Krantz, R. Luce, and A. Tversky, 1989: *Foundations of Measurement*, Vol. II (Academic Press, San Diego).

van Fraassen, B., 1981: 'Probabilistic Semantics Objectified', *Journal of Philosophical Logic* **10**, 371–394, 495–510.

Department of Philosophy,
University of Oklahoma,
Norman OK 73019,
U.S.A.