BARRY M. LOEWER

COTENABILITY AND COUNTERFACTUAL LOGICS

I

There are two accounts of the truth conditions of counterfactual statements which have been developed. According to David Lewis's[1] formulation of the possible world account "if A were the case then B would be the case" is true if there are possible worlds in which A and B are true which are more similar to the actual world than any possible world in which A is true and B is false. The other approach is the metalinguistic theory which counts a counterfactual as true if its antecedent together with certain auxiliary statements and laws of nature implies its consequent. Nelson Goodman[2] has developed this account and it is his version of it which we will investigate in this paper. Lewis has argued that the metalinguistic theory is compatible with his, while others[3] have claimed that the two approaches are fundamentally different. In this paper we seek to clarify the relationship between the two accounts particularly with respect to the logics which they determine.

In order to state the two approaches with precision we will formulate them for a propositional language $\mathcal{L}$ with finitely many atoms to which the binary connective $>$ is added. Lewis semantics for $\mathcal{L}$ are fomulated in terms of the concept of a system of spheres. A system of spheres is a three-tuple $\langle W, w^*, \$\rangle$ where W is a set of possible worlds, $w^* \in W$ is the actual world, and $\$$ is a function which assigns to each $u \in W$ a subset of the power set of W, $P(W)$, which is totally ordered by set inclusion and which has $\{u\}$ as its minimal member. Lewis suggests that we think of the members of $\$(u)$ as forming spheres centered on u. $S \in \$(w)$ is said to be A-permitting if A is true at some $v \in S$ and A-necessitating if A is true at all $v \in S$. If v belongs to a sphere around w to which u does not belong then v is said to be more similar than u to w. It will simplify our discussion to assume that if there is an A-permitting sphere in $\$(w)$ then there is a smallest A-permitting sphere in $\$(w)$.[4] A Lewis model M is a system of spheres and a function I which assigns to each atomic sentence A of $\mathcal{L}$ a subset of W, $\|A\|_I$ (the set of worlds at which A is true). Truth functional compounds are treated in the usual manner. Truth conditions for counterfactuals are given by

Journal of Philosophical Logic 8 (1979) 99–115. 0022–3611/79/0081–0099$01.70

(L) $I(A > B)$ is true at w of $\langle W, w^*, \$\rangle$ (i.e., $w \in \|A > B\|_I$) if and only if there is an A-permitting $S \in \$(w)$ which is $A \rightarrow B$ necessitating or else there is no A-permitting $S \in \$(w)$.

A is valid iff $\|A\|_I = W$ for every Lewis model. We will call the system of logic thus characterized as well as the semantics which characterize it 'Lewis'.

II

Goodman's discussion of counterfactuals is informal and he does not attempt to provide formal semantics. His primary interest seems to be in analyzing counterfactuals in terms of non-counterfactuals. He seeks to explain the truth conditions of counterfactuals in terms of notions like logical implication and law of nature. In the course of his discussion he makes a number of remarks which, we will argue, suggest a certain formulation of a counterfactual logic. Goodman's penultimate analysis of counterfactuals is:

> Our rule thus reads that a counterfactual $A > B$ is true if and only if (i) there is some set D of true sentences such that D is compatible with B and with $-B$, and such that $A \cdot D$ is self-compatible and leads by law to B; (ii) while there is no set H compatible with B and with $-B$, and such that $A \cdot H$ is self-compatible and leads by law to $-B$ (p. 13).[5]

The idea behind this account seems to be this: To determine the truth value of, for example, "If that match were scratched it would light", one tries to find some laws of nature and some true auxiliary conditions D which together with "that match is scratched" imply "that match lights". Such laws and conditions satisfy (i) and this establishes the truth of "If this match were scratched then it might have lit". Satisfaction of requirement (ii) guarantees that "if that match were scratched it might not have lit" is false. Together, satisfaction of (i) and (ii) establish the truth of the counter-factual.

Goodman remarks that his rule "involves a certain redundancy; but simplification is not in point here, for the criterion is still inadequate" (p. 13). The inadequacy is that the rule counts some counterfactuals as true which are false. Goodman's example is: If match m had been scratched, it would not have been dry. His reason is that "match m is scratched" together with "it does not light and it is well made and oxygen is present, etc." law-implies "It was not dry". Also, there would seem to be no suitable set of sentences H such that "match m is scratched" and H leads by law to "match m is dry". According to Goodman "the trouble is caused by including in our D a true statement which though compatible with A would not be true if A were" (p. 15). The offending statement in this example is "it does not light". Goodman suggests that the statements in D (and H) be limited to statements with which A is *cotenable*. He defines "A is cotenable with D, and the conjunction $A \cdot D$ is self-cotenable, if it is not the case that D would not be true if A were" (p. 15).

Since D is a set of sentences, "A is cotenable with D" is ambiguous. It can be interpreted either as "A is cotenable with each member of D" or as "A is cotenable with the conjunction of the members of D". Goodman seems to mean the latter. To resolve the ambiguity in this way, we will define cotenability as a relation between sentences rather than as a relation between a sentence and a set of sentences. When Goodman speaks of A being cotenable with the set D we will interpret him as meaning that A is cotenable with the conjunction of the members of D. The definition is as follows: 'A is cotenable with B, (Cot(A, B)), iff $-(A > -B)$'.

If the sets of statements D, H in Goodman's rule are restricted so that Cot(A, D) and Cot(A, H) then the requirement that D and H are compatible with B and $-B$ is redundant. Taking this into account our proposed formulation of Goodman's truth conditions for counterfactual statements is:

(G) $A > B$ is true iff (i) there is a finite set of true statements D such that Cot(A, D) and A and D together with some laws imply B, and (ii) there is no set of true statements H such that Cot(A, H) and A and H together with some laws imply $-B$.

Goodman remarks that his account faces a serious difficulty. In order to determine whether or not $A > B$ is true one must be able to determine whether or not A is cotenable with D for various D and that involves being able to determine whether or not $(A > -D)$ is true. As he remarks "In other words to establish any counterfactual, it seems that we first have to determine the truth of another. If so, we can never explain a counterfactual except in terms of others, so that the problem of counterfactuals must remain unsolved" (p. 16).

We will argue that despite its failure as an analysis of counterfactuals, Goodman's account does provide a framework for a counterfactual logic. We will show that Goodman's truth conditions are compatible with Lewis's and in fact a logic based on Goodman's account is a subsystem of Lewis's.

III

Goodman confines his discussion to counterfactual statements which do not contain iterated occurrences of $>$. We will first formulate a logic based on rule (G) for languages restricted so as not to contain iterations of $>$. The extension of the logic to all formulas of $\mathcal{L}$ will be straightforward.

To construct a logic around (G) we will need some additional principles which Goodman does not explicitly formulate. After presenting the logic we will argue that in fact Goodman implicitly assumes these principles in the course of his discussion. Some abbreviations will be useful; $\Box A$ for $-A > A$ and $\Diamond A$ for $\mathrm{Cot}(A, A)$. Also, let Δ be the set of laws.

Our formulation is in terms of the notion of a Good set. A set Γ of sentences of $\mathcal{L}$ (restricted to sentences which contain no iterated occurrences of $>$) is Good iff Γ is truth functionally consistent and maximal and satisfies the following conditions:

(1) $\Delta \subseteq \Gamma$

(2) $\Diamond A$ iff $\Delta \cup \{A\}$ is truth functionally consistent.

(3) If $\Box(A \leftrightarrow B) \in \Gamma$ then $(A > C \leftrightarrow B > C) \in \Gamma$.

(4) If $(A > B)\cdot(A > C) \in \Gamma$ then $A > B\cdot C \in \Gamma$.

(5) $A > B \in \Gamma$ iff either $\Box - A \in \Gamma$ or the following two conditions hold: (i) there is a $D \in \Gamma$ such that $\mathrm{Cot}(A, D) \in \Gamma$ and $\Box(A \cdot D \rightarrow B) \in \Gamma$ and (ii) there is no $H \in \Gamma$ such that $\mathrm{Cot}(A, H) \in \Gamma$ and $\Box(A \cdot H \rightarrow - B) \in \Gamma$.

A set of sentences is J consistent iff there is some Good set which contains it. B is a theorem of J, $\vdash_J B$, iff $\{-B\}$ is not J consistent. We extend the system J to a system G suitable for languages which permit iterations of $>$ as follows: The axioms of G are the theorems of J (J is obviously decidable). A G proof is a sequence of formulas each one of which is either an axiom or obtained from previous members of the sequence by substitution, modus ponens, or necessitation. Any formula which occurs in a G proof is a G theorem. We call the system of logic just characterized G.[6]

There are a number of differences between our formulation of Goodman's truth conditions (G) and the corresponding conditions (5). According to (G) $A > B$ is false when A is not self-cotenable but (5) counts such formulas as true. Without this modification $-\mathrm{Cot}(A, A) \rightarrow \mathrm{Cot}(A, A)$ would be valid since according to (G) if $-\mathrm{Cot}(A, A)$ then no counterfactual with antecedent A, inclucing $A > A$, is true. Since logical contradictions are not self-cotenable this means that rule (G) is inconsistent. Evidently the inconsistency escaped Goodman's notice. It would be possible to modify the definition of cotenability and preserve the principle that $A > B$ is false when $-\mathrm{Cot}(A, A)$. We find the modification in (5) preferable since it brings Goodman's system into agreement with Lewis. Another difference between (G) and (5) is that laws are not explicitly mentioned in the latter. However, $\Box A$ is true just in case A is implied by Δ. So $\Box(A \cdot D \rightarrow B)$ says that A and D law-imply B. We will discuss the role of laws in G more fully later.

We have added four conditions, (1)–(4), to our modification of Goodman's truth conditions in constructing the logic G. We will now argue that at least three of these are implicitly assumed by Goodman. Conditions (1) and (2) characterize $\Box$ as lawful necessity. It is straightforward to show in G the $\Box$ is a T necessity operator. Condition (1) in effect says that laws are true. This is certainly also assumed by Goodman. He doesn't explicitly formulate (2). However, prior to introducing the concept of cotenability he does say that a sentence will be considered to be 'self-compatible' just in case it is logically consistent with laws (p. 10). Since in subsequent formulations of the truth conditions of counterfactuals cotenability replaces compatibility, it is not unreasonable to suppose that Goodman would agree to (2).

Evidence that Goodman would accept (3) is to be found in his discussion of the following pair of counterfactuals (p. 15):

> If New York City were in Georgia, then New York City would be in the South. If Georgia included New York City, then Georgia would not be entirely in the South.

He remarks that these two counterfactuals present a problem since both are acceptable yet their antecedents are logically equivalent. Apparently

Goodman believes that it is a consequence of his account that if A and B are logically equivalent then at most one of $A > C, B > -C$ can be true. In fact, once our modification for impossible antecedents is made, it is easy to see that Goodman's truth conditions imply that if A is possible then only one of $A > C, A > -C$ can be true. It seems reasonable to infer that it is because he implicitly assumes that logically equivalent statements can be substituted in counterfactuals that he then finds the pair under discussion problematic. Goodman's solution to the problem is that when we assent to the two counterfactuals we are construing their antecedents in such a way that they are not logically equivalent. He suggests that the first one is understood as having the antecedent "If New York City were in Georgia, and the boundaries of Georgia remained unchanged" and the second as containing the antecedent "If Georgia included New York City, and the boundaries of New York City remained unchanged" (p. 16). Whatever the merits of this resolution, Goodman's discussion of the problem does strongly suggest that he implicitly assumes (3).[7]

Evidence that Goodman's discussion implicitly assumes (4) is more difficult to find. The plausibility of (4) can be made evident by noting that it is equivalent to the condition: if $\mathrm{Cot}(A, B) \in \Gamma$ then either $\mathrm{Cot}(A, B \cdot C) \in \Gamma$ or $\mathrm{Cot}(A, B \cdot -C) \in \Gamma$. It certainly seems that anyone willing to accept "If A were true then B might be true" ought to be willing to accept either "If A were true then $B \cdot C$ might be true" or "If A were true then $B \cdot -C$ might be true". Besides its plausibility our main motivation for including (4) in the definition of 'Good set' is that it facilitates the construction of semantics for G. The system obtained by dropping (4), we will call it G-4, has semantics which are a bit more complicated than semantics for G. We will discuss semantics for both systems in Section V.

It is not difficult to show that the following are theorems of G (and also of G-4).

(6) $A > A$

(7) $\Diamond A \rightarrow ((A > B) \rightarrow \mathrm{Cot}(A, B))$

(8) $(A > B) \rightarrow (A \rightarrow B)$

(9) $(A \cdot B) \rightarrow (A > B)$

(10) $(A > B) \cdot \Box(B \rightarrow C) \rightarrow (A > C)$

(11) $A > B \leftrightarrow (\Box - A \vee ((\mathrm{Cot}(A, A \rightarrow B) \cdot - \mathrm{Cot}(A, A \rightarrow - B))$

(12) $\Box(A \rightarrow B) \rightarrow (A > B)$.

We will prove (8) and (9) by showing that it is not possible to embed their negations in Good sets.

Proof of (8). Suppose that Γ is Good and that $A > B \in \Gamma$ and $(A \cdot - B) \in \Gamma$. By (5) either $\Box - A \in \Gamma$ or conditions (i) and (ii) are satisfied. If the former then $-A \in \Gamma$ so, contrary to the assumption, Γ is not Good. Suppose the latter, that (i) and (ii) hold. Then by (i) there is a $D \in \Gamma$ such that $\mathrm{Cot}(A, D) \in \Gamma$ and $\Box(A \cdot D \rightarrow B) \in \Gamma$.

Since Γ is Good, $B \in \Gamma$. But we already have $A \cdot - B \in \Gamma$ and hence $-B \in \Gamma$ and so contrary to the assumption Γ is not Good. So we cannot embed the negation of (8) in a Good set.

Proof of (9). Suppose that Γ is Good and that $A \cdot B \in \Gamma$ and $-(A > B) \in \Gamma$. We will show that (i) and (ii) are satisfied with respect to $A > B$ and so $A > B \in \Gamma$. Clearly there can be no $H \in \Gamma$ such that $\mathrm{Cot}(A, H) \in \Gamma$ and $A \cdot H$ implies $-B$ since if there were then $-B$ would belong to Γ contrary to the assumption that Γ is Good. Hence, (ii) is satisfied. To show that (i) holds we let D be $A \rightarrow B$. $A \rightarrow B \in \Gamma$(since $A \cdot B \in \Gamma$) and with A implies B. So all we need to show is that $\mathrm{Cot}(A, A \rightarrow B) \in \Gamma$. If $-\mathrm{Cot}(A, A \rightarrow B) \in \Gamma$ then $A > -(A \rightarrow B) \in \Gamma$. By (8) $A \rightarrow (-A \cdot B) \in \Gamma$. But then $-A \in \Gamma$. Since $A \in \Gamma$, this contradicts the assumption that Γ is Good. We conclude that $\mathrm{Cot}(A, A \rightarrow B) \in \Gamma$. But then (i) and (ii) both hold and so $A > B \in \Gamma$.

It may be surprising to discover that (9) is a theorem of G. Jonathan Bennett[8] suggests that the fact that (9) is valid in Lewis is a defect in his account and claims that it is an advantage of the metalinguistic approach that (9) is not valid in it. We do not wish to argue over the merits of (9) here. In any case, both G and Lewis require only slight modifications to render (9) invalid. This is accomplished in Lewis by replacing the condition that $\{u\} \in \$(u)$ with the condition that $u \in S$ for each $S \in \$(u)$ and it is accomplished in G by dropping the requirement in (ii) that $H \in \Gamma$. I am uncertain as to whether or not Goodman intended that the statements in H in (ii) should be restricted to true statements (which is what the requirement that they belong to Γ comes to). It is likely that had he realized that the validity of (9) is the result he would have dropped the restriction.

While (7)–(12) are theorems of both Lewis and G the following are not theorems of either system:

(13) $(A > B) \rightarrow (A \cdot C > B)$

(14) $(A > B) \rightarrow (-B > -A)$

(15) $((A > B) \cdot (B > C)) \rightarrow (A > C)$

(16) $(A > B) \vee (A > -B)$.

The non-theoremhood of (13)–(15) is characteristic of counterfactual logics. It is interesting to note that had we construed Goodman's definition of cotenability as requiring that A be cotenable with each member of D rather than with the conjuction of the members of D we would have obtained a system in which (16) is valid.

The following are theorems of Lewis but not of G:

(17) $((A \vee B) > C) \rightarrow ((A > C) \vee (B > C))$

(18) $((A > C) \cdot (B > C)) \rightarrow ((A \vee B) > C)$

(19) $(A > C) \cdot \mathrm{Cot}(A, A \cdot B) \rightarrow (A \cdot B > C)$.

The significance of these formulas will become apparent later in our discussion.

IV

The most striking difference between Lewis's and Goodman's accounts is the roles that each assigns to scientific laws. In Goodman's, but not in Lewis's laws enter into the formulation of truth conditions for counterfactuals. According to (G) $A > B$ is true only if there are auxiliary conditions D such that $\mathrm{Cot}(A, D)$ and laws L such that $A \cdot D \cdot L$ implies B. Goodman seems to intend that these laws should be essential to the implication. However, in view of the fact that if $A > B$ then $\mathrm{Cot}(A, A \rightarrow B)$ it does not seem possible to capture the idea that laws are essential for going from A to B. In our formulation (5) we require that if $\mathrm{Cot}(A, A)$ then $A > B$ is true only if there is a D such that $\mathrm{Cot}(A, D)$ and $\Box(A \cdot D \rightarrow B)$. Since $\Box$ is law-necessity, laws do enter into the truth conditions of $A > B$. But since $\mathrm{Cot}(A, A \rightarrow B)$ when $\Diamond A$ and $A > B$, laws cannot be *essential* to the implication of B.

One may feel that we have not captured in G the full role that laws are intended to play in Goodman's analysis. However, it seems to us that the special status assigned to laws by Goodman is epistemological. It may very well be that the usual way in which we argue for the truth of $A > B$ is by deriving B from A, cotenable auxiliary statements, and laws. And we may sometimes argue against Cot(A, D) by showing that $A \cdot D$ is inconsistent with some laws. Since on our account laws are necessarily true they can play these epistemological roles.

Lewis does not agree that laws are cotenable with every self-cotenable statement. This is an important metaphysical and epistemological difference between Lewis's and Goodman's accounts. However, this difference does not show up in their logics since we have placed no conditions on Δ; it may even be empty. For this reason we will ignore laws in our discussion of the two logics.

Lewis argues that his truth conditions and the metalinguistic truth conditions are compatible. He defines "A is cotenable with D" as "there exists an $S \in \$(w^*)$ which is A-permitting and D-necessitating" (p. 57). Using this definition of cotenability it is easy to show that $A > B$ is true at w^* of a Lewis model if and only if (5) is satisfied with respect to w^* of the model. Lewis claims that this establishes the compatibility of the two accounts. However, it is clear that Lewis's definition of cotenability is much stronger than Goodman's. Bennett also notices that Lewis's construal of cotenability differs from Goodman's and says of Lewis's argument that the possible world and the metalinguistic accounts are compatible: "The derivation, which at first sight looks like a rather dashing capture of enemy territory is really no such thing. It goes through on Definition 1 (Definition 1 is Lewis's definition) . . . of cotenability but fails for Definition 3 (Goodman's definition), and so it fails absolutely". However, it is not difficult to show that even if we use Goodman's definition, Lewis's truth conditions and our formulation of Goodman's truth conditions by (5) are compatible.[9] We can do this by showing that $A > B$ is true at w^* of the model $\langle\langle W, w^*, \$\rangle I\rangle$ iff the conditions in (5) are satisfied at w^*. If $A > B$ is true at w^* then either there is no A-permitting sphere in $\$(w^*)$ which is $A \rightarrow B$-necessitating or there is an A-permitting sphere in $\$(w^*)$ which is $A \rightarrow B$-necessitating. In the first case $-$ Cot(A, A) so (5) is satisfied. In the second case (i) there is a true D, namely $A \rightarrow B$, such that Cot(A, $A \rightarrow B$) and $\Box(A \cdot (A \rightarrow B) \rightarrow B)$. Also (ii) there can be no true H such that Cot(A, H) and $\Box(A \cdot H \rightarrow - B)$. For if there were such an H then no sphere centered on w^* could be A-permitting and $A \rightarrow B$-necessitating since it would contain a world at which A is true and B is false. So if $A > B$ is true at w^* then (5) is satisfied at w^*. To demonstrate the converse suppose that conditions (5) are satisfied with respect to $A > B$ at w^*. Then either $-\operatorname{Cot}(A, A)$ or (i) and (ii) hold. If the first then $\$(w^*)$ contains no sphere which is A permitting so $A > B$ is true. If (i) holds then there is a sphere in $\$(w^*)$ which is A-permitting and if (ii) holds then there is an A-permitting sphere which is $A \rightarrow B$-necessitating. So Lewis's truth conditions hold.

V

The preceding argument shows that within Lewis's semantical framework his and Goodman's truth conditions (as formulated by (5)) coincide. While this shows that the two accounts are compatible it is still of interest to compare the two counterfactual logics G and Lewis.

To make the comparison more perspicuous we will work with a formulation of Lewis's semantics in terms of selection functions. A selection function structure is a triple $\langle W, w^*; f\rangle$ where W is a set of worlds, w^* is the actual world, and f is a binary function from $WXP(W)$ to $P(W)$. A Lewis selection function model is a selection function struture and valuation I which satisfies these conditions (Lewis, p. 58).

(a) if $w \in \|A\|_I$ then $f(w, A) = \{w\}$ (we write $f(w, A)$ for $f(w, \|A\|_I)$)

(b) $f(w, A) \subseteq \|A\|_I$

(c) if $\|A\|_I \subseteq \|B\|_I$ and $f(w, A) \neq \wedge$ then $f(w, B) \neq \wedge$ ($\wedge$ is the empty set)

(d) if $\|B\|_I \subseteq \|A\|_I$ and $\|B\|_I \subseteq f(w, A)$ then $f(w, B) = f(w, A) \cap \|B\|_I$.

Truth conditions for $A > B$ are given by

L' $I(A > B, w)$ is true iff $f(w, A) \subseteq \|B\|_I$.

The remaining conditions on I and the definitions of consistency and validity are as expected.

Starting with a system of spheres (satisfying the limit assumption) a selection function structure can be constructed by keeping W and w^* the same and defining $f(w, A)$ as the intersecton of $\|A\|_I$ and the smallest A-permitting sphere centered on w. The two structures are equivalent in that if I is an interpretation defined for the system of spheres and I' is defined for the corresponding selection function structure and I and I'

agree on atomic sentences then I and I' agree on all sentences. It is also possible to start with a selection function structure and define an equivalent system of spheres (Lewis, p. 58). So the two semantics determine the same logic.

It is not difficult to show that if A is G-valid then A is Lewis-valid. In fact the logic G is complete with respect to the selection function semantics from which condition (d) is omitted. We will call the system based on these semantics L^-. The proof that G is complete with respect to the semantics L^- is straightforward. In fact G less the instances of (9) is equivalent to a system investigated by Brian Chellas which he calls $CK + ID + MP$.[10] Chellas proves that this system is complete with respect to semantics just like L^- except that (9) is replaced by: if $w \in \|A\|_I$ then $w \in f(w, A)$.

In Section III we briefly discussed the logic G-4 that results from dropping (4) from the definition of a Good set. In G-4 $(A > B)\cdot(A > C) \rightarrow (A > B\cdot C)$ is not a valid schema. This makes it impossible to construct selection function semantics for G-4. However, semantics for G-4 can be provided by using Scott and Montague's neighborhood semantics.[11] An N-structure is a triple $\langle W, w^*, F\rangle$ where W and w^* are as before and F is a function from $WXP(W)$ to $PP(W)$. F assigns to each world w and proposition ϕ the set of propositions which are cotenable with ϕ at w. An N-model is an N-structure and interpretation V which satisfies the following conditions:

(j) not $\|-A\|_V \in F(w, A)$

(k) if $w \in \|A\|_V$ then $\{w\} \in F(w, A)$

(l) if $\|B\|_V \in F(w, A)$ and $\|B\|_V \subseteq \|C\|_V$ then $\|C\|_V \in F(w, A)$

(m) if $\|A\|_V \subseteq \|B\|_V$ and $\Lambda = F(w, B)$ then $\Lambda = F(w, A)$.

Truth conditions for counterfactuals are given by

> $V(A > B)$ is true at w of $\langle W, w^*, F\rangle$ if and only if either $F(w, A) = \Lambda$ or $\|B\|_V \in F(w, A)$ and not $\|-B\|_V \in F(w, A)$.

The logic G-4 is complete with respect to these semantics. If we add the condition

(n) if $\|A\|_V \in F(w, A)$ then either $\|A\| \cap \|B\|_V \in F(w, A)$ or $\|A\|_V \cap \|-B\|_V \in F(w, A)$

then the resulting semantics is equivalent to L^-. This can be seen by noting that if we define a selection function f as $u \in f(w, A)$ iff $\{u\} \in F(w, A)$ then f satisfies (a), (b), and (c) and the N semantics truth condition reduces to L'.

VI

Lewis's semantics are obtained from L^- by requiring that selection functions satisfy (d). It is (d) that results in an ordering of possible worlds as more or less similar to the actual world. This can be seen as follows: We define wRu as $f\{w, u\} = \{w\}$ or $f\{w, u\} = \{w, u\}$. (We take w^* to be the first argument place of f and suppress reference to it for the sake of typographical clarity.) It can be proved that R is a weak ordering on W, that w^* is the unique minimal member of W, and that $f(A)$ = the set of R-least members of $\|A\|$. So Lewis's account goes beyond our formalization of Goodman precisely in requiring a ranking of possible worlds as more or less *similar* to the actual world. Exactly what is involved in saying that R is a similarity relation will be discussed at the conclusion of this paper.

The question naturally arises as to whether or not our reasoning with counterfactuals actually does involve ranking possible worlds. Notice that adding (d) to L^- is equivalent to adding to G the instances of (19) as theorems. An inference corresponding to (19) is:

(20) If a liberal had been elected then we would not be in the mess we are in now.
If a liberal had been elected then a northeastern liberal might have been elected.

If a northeastern liberal had been elected then we would not be in the mess we are in now.

This argument certainly appears to be valid. Insofar as we have reason to believe that every argument of this form is valid we have reason to adopt (d).

John Pollock[12] claims to have discovered the following counterexample to (19). Let S, T, U be the statements, "My car is painted black", "My garbage can blew over", and "My maple tree died". We suppose that S, T, and U are false and unrelated. Pollock considers the substitution instance of (19)

(21) $((S \vee T) > -U) \cdot \mathrm{Cot}(S \vee T, U \vee T) \rightarrow ((S \vee T) \cdot (U \vee T) > -U)$

Formula 21 implies

(22) $\mathrm{Cot}((S \vee T), \neg S) \cdot \mathrm{Cot}((S \cdot U) \vee T, \neg T) \rightarrow \mathrm{Cot}(S \vee T, U)$.

Pollock argues that $\mathrm{Cot}(S \vee T, U)$ is false. His reason is that since the color of my car and the state of my garbage can are unrelated to the condition of my tree, even if $S \vee T$ were true my tree would not die. Pollock claims that the antecedent of (22) is true since "disjunctions whose disjuncts are unrelated to one another cannot necessitate either disjunct".[13] So both $\mathrm{Cot}(S \vee T, \neg S)$ and $\mathrm{Cot}(S \cdot U \vee T, \neg T)$ are true.

It is difficult to evaluate Pollock's alleged counterexample since the English version of (22) is rather complicated. The difficulty is exacerbated by the fact that the English counterfactual statements whose truth values we are attempting to assess have disjunctive antecedents. These are known to create special problems. A number of authors have argued that English counterfactual statements with disjunctive antecedents should not be paraphrased by formal language counterfactuals with disjunctive antecedents.[14] Another difficulty with the example is that the claim that disjunctions whose disjuncts are unrelated to one another cannot necessitate either disjunct, which Pollock appeals to in order to support the truth of the antecedent of (22), is not clearly true. Consider the apparently unrelated and false statements "Nixon was impeached", "Attila the Hun is still alive". The counter factual "If either Attila the Hun were still alive or Nixon had been impeached then Nixon would have been impeached" seems to be true. One might reason that it is true since, although the disjuncts are unrelated, Attila the Hun's still being alive is so far fetched that if the disjunction were true it would have to be because Nixon were impeached. However, we do not put too much reliance on this objection since, as we noted, counterfactuals of this kind are difficult to evaluate.

Even if we accept Pollock's counterexample it can be argued that our reasoning with counterfactuals involves similarity comparisons among possible worlds. If we add the instances of (17) and (18) as theorems to G then semantics for the resulting system, G^*, also involves ordering possible worlds. Schemes (17) and (18) correspond to the semantic conditions

(e) $f(A) \subseteq f(A \vee B)$ or $f(B) \subseteq f(A \vee B)$

(g) $f(A \vee B) \subseteq f(A) \cup f(B)$.

Although (d) implies both (e) and (g) within L^- the converse does not hold. However, if we define wR^*u as true iff $f\{w, u\} = \{w\}$ then it can be shown that R^* partially orders W and w^* is the unique first R^*-element.[15] Furthermore, $f(A)$ is the set of R^*-least members of $\|A\|$. So it seems that by accepting (17) and (18) as valid principles, and they certainly seem to be valid, one is committed to ranking possible worlds as more or less similar to the actual world.

The system G^* is an interesting one since it is immune to Pollock's alleged counterexamples, it does not contain the dubious (16), and yet it still involves ranking possible worlds.[16] However, arguments like (20) present a problem for G^*. Although (20) seems to be valid the natural paraphrase of (20) is not G^* valid.

VII

We conclude with a brief discussion of what is involved in saying that R (or R^*) is a similarity relation. Since R is a ranking of worlds with w^* as its first member it is not unnatural to say that if uRw then u is more similar to w^* than is w. In a similar vein one might say that since distance from Stanford orders institutions of higher learning, U.C.L.A. is more similar to Stanford than is Princeton. Of course, similarity in this last statement must be understood as a similarity with respect to distance from Stanford. But how is 'similarity' to be understood in Lewis's semantics? Lewis himself wavers between two answers to this question. He sometimes claims that the notion of similarity involved is an antecedently understood notion of 'overall similarity' (Lewis, p. 92). At other times his view seems to be that the similarity relation is a new technical concept introduced to provide semantics for counterfactuals. Referring to the relation R he remarks "Not for nothing did I call it primitive".[17] There is an important difference between the two positions. If R is taken to be an antecedently understood relation of overall similarity then Lewis's account is an *analysis* of counterfactuals in terms of R. If R is a primitive technical concept then Lewis cannot claim to have provided an analysis of the truth conditions of counterfactuals but only a formalization of semantics for a counterfactual logic.

In some places Lewis quite clearly takes the position that his account is an analysis of counterfactuals in terms of the vague but antecedently understood notion of overall similarity. At a number of points his arguments for and against certain principles of counterfactual logic appeal to supposed features of overall similarity. One argument he gives for the validity of (9) depends on the claim that no world is as similar to the actual world as the actual world is to itself. Another example is Lewis's argument against the semantic principle that for each possible A there is a least A-permitting sphere (p. 19). He claims that for every world w which contains a printed line on page 20 of *Counterfactuals* (say the length of the line in w is $1 + \xi''$) there is a world u which is more similar to the actual world (the length of the line in u is $1 + \xi''/2$). These arguments have force only if the notion of similarity appealed to is antecedently understood.

If Lewis is providing an analysis of counterfactuals in terms of overall similarity then his account is open to certain objections. Kit Fine remarks[18] that on Lewis's analysis "If Oswald had not shot Kennedy then someone else would have" is true, since a world in which Oswald did not shoot Kennedy but someone else did is, overall, more similar to the actual world than any world in which no one shot Kennedy. Fine bases his judgment of overall similarity on the observation that had Kennedy served out his term the changes in world history would have been much more profound than the changes required to accommodate another assassin. If Fine is correct in his assessment of overall similarity and in thinking that the counterfactual at issue is false, then we have a clear counterexample to Lewis's analysis. Of course, Lewis might dispute Fine's evaluation of overall similarity, arguing that the falsity of the counterfactual shows Fine's assessment to be mistaken. The dispute seems impossible to settle, however, since the concept of overall similarity to which Lewis appeals is not well enough understood to support an analysis of counterfactuals.

In this paper we have shown that Lewis's semantics for counterfactuals formulated in terms of similarity comparisons among possible worlds and Goodman's analysis formulated in terms of cotenability are compatible. We extracted from Goodman's tentative and sketchy account a counterfactual logic G which, given a few modifications, was shown to be a subsystem of Lewis. Lewis's account goes beyond Goodman's precisely in requiring similarity comparisons among possible worlds. Finally, we argued that, although our reasoning with counterfactuals does involve a similarity ordering of worlds, the concept of similarity is primitive and does not support an analysis of counterfactuals.

University of South Carolina

NOTES

* I would like to thank Mark Leach, Professors James Oliver, Kit Fine, Donald Nute, Thomas McKay, and especially the referees of the *Journal of Philosophical Logic* for criticisms of earlier versions of this paper. This paper was written while the author held a Summer research grant from the University of South Carolina.

1 David Lewis, *Counterfactuals*, Oxford, 1973. Parenthetical page references to Lewis are to this book.

2 Nelson Goodman, *Fact, Fiction and Forecast*, 2nd edn., The Bobbs-Merrill Company, Inc., 1965. Parenthetical page references to Goodman are to this book.

3 Especially see: Kit Fine, 'Critical Notice of Counterfactuals', *Mind*, Vol. LXXXIV No. 335, pp. 451–458, and Jonathan Bennett, 'Counterfactuals and Possible Worlds', *Canadian Journal of Philosophy* 4, 381–402.

4 Lewis calls this condition 'the Limit Assumption' (Lewis, p. 20). Lewis claims that its violation is required for the correct interpretation of, for example, "If this stone were heavier than 10 lb I would not have been able to lift it". There is no smallest sphere which permits "This stone is heavier than 10 lb", since there is no smallest weight greater than 10 lb. However, it turns out that the set of valid formulas of Lewis is the same whether or not this assumption holds. See David Lewis, 'Counterfactuals and Comparative Possibility', *Journal of Philosophical Logic* (1973), 418–446.

5 We took some liberties in reproducing this passage from Goodman to avoid conflict with Lewis's notation. In a footnote (p. 13) Goodman adds another condition to his rule in order to overcome an objection raised by W. T. Parry. Parry pointed out that no counterfactual satisfies the rule as stated for one can always take $A \supset -B$ as H. To deal with this Goodman adds the requirement that D and H should not follow by law from $-A$. In the final version of the rule Goodman further restricts D, H, to statements with which A is cotenable. This last restriction makes the requirement introduced to answer Parry's objection unnecessary.

6 The characterization of G is somewhat peculiar since it mixes model set and ordinary proof methods. The mixture is required to stay close to Goodman's formulation while extending it to languages with iterated occurrences of $>$. It is worth noting that none of the theorem schemas of G involve iterations of $>$. The same holds for Lewis. So the extension to languages which allow for iterations really introduces no *new* logical principles which govern $>$.

7 If we drop (3) the resulting logic is still recognizably a conditional logic. However, semantics for this logic are somewhat nonstandard since the worlds one looks at to evaluate $A > B$ will depend on the syntactic form of A not on the proposition expressed by A. Donald Nute has developed semantics of this sort. Nute, 'Simplification and Substitution of Counterfactual Antecedents', *Philosophia*, forthcoming.

8 Bennett, *op. cit.* p. 391.

9 *Ibid*, p. 391. Our criticism of Bennett is somewhat unfair since Bennett seems to be using a version of the metalinguistic account which contains (i) but not (ii). It is true

that one cannot derive Lewis's truth conditions from (i) alone. Bennett does not explicitly formulate a version of the metalinguistic account.

[10] Brian Chellas, 'Basic Conditional Logic', *Journal of Philosophical Logic* (1975), 135–153.

[11] For a discussion of neighborhood semantics, see Segerberg K., An Essay in Classical Modal Logic, *Filosofiska studier utgivna av Filosofiska Foreningen och Filosofiska Institutionen vid Uppsala* Universitet nr 13, Uppsala, 1971. Chellas, *op. cit.* p. 144, also constructs neighborhood semantics for a conditional logic in which $(A > B \cdot A > C) \rightarrow (A > B \cdot C)$ is not valid.

[12] John Pollock, *Subjunctive Reasoning*, D, Reidel, Dordrecht, 1976, pp. 43–44.

[13] *Ibid*, p. 44.

[14] Barry Loewer, 'Counterfactuals with Disjunctive Antecendents', *Journal of Philosophy*, Vol. LXXIII (1976).

[15] A proof that R^* is transitive follows: Suppose wR^*u and uR^*v. So $f\{w, u\} = \{w\}$ and $f\{u, v\} = \{v\}$. by (g) $f\{w, u, v\} \subseteq f\{w, u\} \cup f\{v\}$ and $f\{w, u, v\} \subseteq f\{w\} \subseteq f\{u, v\}$. This implies that $f\{w, u, v\} \subseteq \{w\}$ and since $f(A) = \Lambda$ we have shown that $f\{w, u, v\} = \{w\}$. By (e) $f\{w, u, v\}$. But this implies that $f\{w, v\} = \{w\}$ and so wR^*v. R^* cannot be extended to a weak ordering in this case since wQu, defined as $f\{w, u\}$, need not be transitive. However, given (d) it can be proved that wQu is transitive and so wRu defined as wR^*u or wQu is a weak ordering.

[16] System G^* is equivalent to Pollock's system *SS*. Pollock, *op. cit.*, p. 42.

[17] David Lewis, 'Causation', *Journal of Philosophy* LXX (1973), p. 559.

[18] Kit Fine, *op. cit.*, p. 452.