Notre Dame Journal of Formal Logic
Volume 21, Number 4, October 1980

An Analysis of the Subjunctive Conditional

CHARLES B. DANIELS and JAMES B. FREEMAN

In this paper we present an analysis of the subjunctive conditional.* Included are both a semantics for and a complete axiomatization of a system which consists of the propositional calculus extended by the addition of an operator symbol representing the subjunctive conditional. We prove that this system is equivalent to the system T of modal logic [9]. We also compare our system with previous analyses of the subjunctive conditional, both of the consequentialist or metalinguistic approach and of the possible worlds *cum* similarity variety. We argue that our system preserves the best of both treatments. In particular, it involves a notion of similarity among possible worlds without incurring the problems Lewis has encountered or incorporating counterintuitive theorems. We also consider certain subjunctive fallacies. It becomes clear under what special conditions transitivity and contraposition do hold for the subjunctive conditional and why under other conditions they fail.

1 The subjunctive conditional A subjunctive conditional is a sentence of the form 'If it were the case that P, then it would be the case that Q'. We follow Stalnaker in symbolizing the subjunctive conditional by '>'. We depart from Stalnaker in that for us '>' is a three-place sentential operator, not a two-place

*This work was done with the aid of Canada Council Grants S74-0551-S1 and S76-0824. We wish to thank Professors R. E. Jennings, J. Michael Dunn, Romane Clark, John Barker, and W. K. Warmbrōd for discussing the problems of the subjunctive conditional with us, and Professors Dunn and Clark especially for making helpful comments on earlier versions of this paper. This work was also improved by presentation at the 1977 Pacific Division APA meetings. We here thank our commentator, Professor Robert L. Martin, and the members of the audience, especially Professor Kit Fine. Thanks are also due to an assessor of Canada Council Project S76-0824 and to Mr. Bryan Campbell.

Submitted November 2, 1977; revised June 25, 1979

sentential operator. Where 'P', 'Q', and 'R' are wffs, so is '$(RP > Q)$'. '$(RP > Q)$' can be read 'If it were the case that P, then it would be the case that Q'. The reader is doubtless left wondering what happened to 'R'. On our view, as on the consequence view presented in [4], [8], [14], and discussed in [2], when a speaker utters a subjunctive conditional, he assumes, if only tacitly, the truth of certain propositions. Consider this example from the Velikovsky controversy: "If the world were to suddenly cease rotating on its axis, then all the objects on its surface would be hurled into outer space". What is assumed here is that the laws of inertial physics would remain the same in a world where such a catastrophe occurred. The speaker intends this statement to hold true where the laws of physics (and perhaps other propositions) true of the actual world also hold. He is not concerned with worlds where the laws of inertial physics true of the actual world cease to hold.

Consider another example: "If a full-scale nuclear war were to occur, then the survivors would be able to procreate only monsters". Here again there is a tacit assumption concerning the laws of genetics and the effect of radiation on humans. The speaker is not interested in worlds where both a nuclear holocaust occurs and the laws of biophysics are suddenly radically altered.

The 'R' in '$(RP > Q)$' represents the conjunction of all the propositions a speaker tacitly assumes when uttering the statement 'If it were the case that P, then it would be the case that Q'. Semantically '$(RP > Q)$' says '"R" is true, and in all accessible worlds, wherever both "R" and "P" are true (i.e., wherever these worlds are sufficiently similar to have "R" true, if "P" is true), then "Q" is true'.

It might be objected here that the 'R' that represents what a speaker assumes when uttering 'If P were the case, then Q would be the case' might itself contain subjunctive conditionals and that this opens our account to a charge of circularity. We certainly admit that 'R' might itself contain or imply propositions of the form '$(AB > C)$'. Indeed, on the account we shall give, where 'A' is any tautology, the sentence 'A' in '$(AA > A)$' implies '$(AA > A)$'. If this is circularity, our system is blessed with it; but we know of no reason to think that its occurrence is more pernicious for us than it is in material implication, where the 'A' in '$(A \supset B)$' may itself contain or imply a sentence '$(C \supset D)$' and the '$(A \supset A)$' in '$((A \supset A) \supset (A \supset A))$' implies the very sentence that contains it.

Notice that different statements may be substituted for 'R' in '$(RP > Q)$'. In making different conditional judgments, we may entertain different tacit assumptions. Furthermore, two speakers may differ in the assumptions they make when asserting counterfactuals with the same antecedent. In fact, it seems not impossible that the same speaker might entertain different implicit assumptions when asserting conditionals with the same antecedent (and perhaps even with the same consequent). An assumption relevant to one situation might be irrelevant to another. The beliefs held one day need not be the beliefs held the next. Indeed, just as the referent of the demonstrative 'that' may change during the course of a sentence, e.g., 'I want that, and I want that, and I want that', so the assumptions behind the utterance of a sentence containing a number of subjunctive 'if . . ., then . . .'s may change along the way. In 'If you were to flip that switch, the iron would heat up, but if there were a big power failure, nothing would work', it is assumed in the first subjunctive clause that there is no power failure; this assumption is not made in connection with the second subjunctive clause.

One might urge that a speaker could assert 'If P were the case, then Q would be the case' outright, no matter how wildly any worlds might vary from the actual world. We can handle this case too, for we see it as a limiting case. Surely the only worlds we need consider are worlds where the laws of logic hold. For 'R' then we may substitute some tautology.

We do not rule out the further case where 'RP' may entail a contradiction. By asserting 'If P were the case, then Q would be the case', where a statement 'R' contradicting 'P' is tacitly assumed, a speaker may simply be expressing his belief in '$\sim P$', especially if 'Q' is absurd in some obvious sense. Below, when we compare our system with the consequence view, we shall discuss why we feel this ternary approach with a variable R is a distinct advantage.

Our claim that subjunctive conditionals have silent components may cause some misgivings. One might query whether we have in mind all the laws operating at a given time when asserting a subjunctive conditional.[1] In the Velikovsky example, does a person making this statement really know all the (relevant) laws of inertial physics? More mundanely, one may say "If I strike this match, it will light", in general ignorance of the laws of physics connecting striking the match and its lighting. How can these be part of the tacit assumption?

But we are not claiming that when one asserts a subjunctive conditional he has all the relevant laws in mind, tacitly assuming them. We claim only that one has something in mind. We may be acquainted in some way with how the world works. On occasion, this may involve only general acquaintance with how scientists say the world works.[2] Our tacit assumptions are drawn from our stock of beliefs about the world. By saying that R symbolizes the tacit assumptions we make, we assert no more than this and not that these tacit assumptions include all the relevant scientific laws or even anything approaching the statement of a law.

Some may find the extreme variability of the statement "If P were the case, Q would be the case" ($P > Q$) bothersome. On our account, if someone asserts ($P > Q$), we really do not know what proposition he has asserted. If two persons assert ($P > Q$), do they really agree? If one asserts ($P > Q$) and the other ($P > \sim Q$) (or more strongly "No, if P were the case, it needn't be the case that Q"), do they really disagree?[3] Unless we know the contextual assumptions, we cannot answer these questions. Yet it seems intuitive that in the former case the two speakers agree, while in the latter they disagree.

We do not see this as a genuine problem for our theory. In the case of indexicals in ordinary language, the speaker's intentions are not spelled out by what he says. To understand what proposition someone asserts by "He did", we must know something further about the speaker's intentions. Below we shall discuss in some detail how on most accounts of the subjunctive conditional, further information plays a role. If we appeal to a speaker's tacit assumptions about the referents of "he" and "did" in understanding "He did", it may be plausible to refer to his tacit assumptions about the world in understanding ($P > Q$).[4]

2 Axiomatics The wffs of the system S are formed in the usual way from the primitive symbols of any version of the propositional calculus augmented by the addition of the three-place sentence operator '$>$'. For convenience we shall suppose that S contains '$\supset$', '$\sim$', and '$\wedge$', understood as usual. We introduce a sentential constant 't' as an abbreviation of an arbitrary tautology of the propositional calculus. We indicate that A is a theorem of S by writing $\vdash A$. The theorems of S consist of all tautologies of the propositional calculus, all instances of the following axiom schemes:

Ax>1 $((AB > C) \supset A)$
Ax>2 $((AB > C) \supset ((A \wedge B) \supset C))$
Ax>3 $((A \wedge \sim(AB > C) \wedge (E_1F_1 > G_1) \wedge \ldots \wedge (E_nF_n > G_n)) \supset \sim(tt > \sim(A \wedge B \wedge \sim C \wedge ((E_1 \wedge F_1) \supset G_1) \wedge \ldots \wedge ((E_n \wedge F_n) \supset G_n))))$,

and all wffs that follow from the foregoing by

R1. If $\vdash A$ and $\vdash (A \supset B)$, then $\vdash B$.
R2. If $\vdash A$, then $\vdash (tt > A)$.

Needless to say, Ax>3 and R2 are anything but intuitive. We shall explain and motivate these in Section 4. Until Section 4, however, we shall pique the reader's curiosity by silence.

3 Semantics We adopt the terminology of [9]. An S-model (subjunctive-conditional-model) is a triple $\langle W, R, V\rangle$, where W is a nonempty set of worlds, R is a binary reflexive relation on W, and V is a valuation satisfying the conditions in [9] for atomic wffs and truth-functional connectives, and satisfying the condition $[V_>]$:

$[V_>]$ For any wffs A, B, and C, and for any $w \in W$, $V((AB > C), w) = 1$ if $V(A, w) = 1$ and for every $w' \in W$ such that wRw', if $V(A, w') = V(B, w') = 1$, then $V(C, w') = 1$; otherwise $V((AB > C), w) = 0$.[5]

Soundness Theorem *If $\vdash A$, then A is S-valid, in symbols $\Vdash A$; i.e., if $\vdash A$, then for every S-model $\langle W, R, V\rangle$ and for every $w \in W$, $V(A, w) = 1$.*

Proof: That the tautologies of the propositional calculus and the $>$-axioms are valid is easily shown. It is also easy to show that R1 and R2 preserve validity.

The notions of a *consistent* wff, a *consistent* set of wffs, and a *maximal consistent* set of wffs we adopt from [9]. Given a consistent set H of wffs of S, we may construct a maximal consistent extension J, as shown in [9], pp. 151-152.

Lemma *If $\{A, \sim(AB > C), (E_1F_1 > G_1), \ldots, (E_nF_n > G_n)\}$ is consistent, then so is $\{A, B, \sim C, ((E_1 \wedge F_1) \supset G_1), \ldots, ((E_n \wedge F_n) \supset G_n)\}$.*

Proof: The result is an easy consequence of R2 and Ax>3.

Completeness Theorem *If A is a consistent wff, there is an S-model in which A is true at some world.*

To show completeness, we shall, given a consistent wff A, construct a system K of maximal consistent sets $K_0, \ldots, K_i, \ldots$. We imitate the procedure of [9], pp. 155-158. K_0 is the result of expanding $\{A\}$ to a maximal consistent set. Having constructed K_i, for every wff $\sim(EF > G)$ in K_i such that $E \in K_i$, we construct a subordinate K_j by taking K_j^0 to be the set whose members are E, F, $\sim G$, and $((P \wedge Q) \supset R)$ for every wff $(PQ > R)$ in K_i. Notice that by the above lemma K_j^0 is consistent. K_j is formed by extending K_j^0 to a maximal consistent set.

We now construct an S-model $\langle W, R, V \rangle$ by letting $W = K$ and wRw' if $w = w'$ or w' is a subordinate of w. For each propositional variable p and every $w \in W$, $V(p, w) = 1$ or 0 according as $p \in w$ or not. For the sentence operators V complies with the standard conditions for S-assignment.

Lemma *Given W, R, and V as defined above, for every wff A of S and for every world $w \in W$, $V(A, w) = 1$ or 0 according as $A \in w$ or not.*

Proof: We proceed by induction on the length of wff. The proof for variables and for truth-functional connectives proceeds as usual. There remain wffs of the form $(AB > C)$.

(a) Suppose that $(AB > C) \in w$. Then by Ax>1, $A \in w$ and $V(A, w) = 1$ by the induction hypothesis. By Ax>2, $((A \wedge B) \supset C) \in w$. Hence if $A \in w$ and $B \in w$, then $C \in w$. Since $(AB > C) \in w$, for all w' subordinate to w, $((A \wedge B) \supset C) \in w'$. Similarly if $A \in w'$ and $B \in w'$, then $C \in w'$. Thus by the definition of R and the induction hypothesis, for all w' such that wRw', if $V(A, w') = 1$ and $V(B, w') = 1$, then $V(C, w') = 1$. By $[V_>]$, $V((AB > C), w) = 1$.

(b) Suppose that $(AB > C) \notin w$. If $A \notin w$, then by the induction hypothesis $V(A, w) \neq 1$, and by $[V_>]$, $V((AB > C), w) \neq 1$. Suppose on the contrary that $A \in w$. By the maximal consistency of w, $\sim(AB > C) \in w$. By the construction of W, there is a subordinate w' such that $\{A, B, \sim C\} \subseteq w'$. By the induction hypothesis and the maximal consistency of w', $V(A, w') = V(B, w') = 1$, and $V(C, w') \neq 1$. By the definition of R, wRw', and by $[V_>]$, $V((AB > C), w) \neq 1$.

This establishes the Completeness Theorem.

Corollary *If $\Vdash A$, then $\vdash A$.*

4 The equivalence of S and T Let us introduce the '$\Box$' of necessity into S by definition:

$$\text{`}\Box A\text{'} =_{df} \text{`}(tt > A)\text{'}.$$

Notice that by $[V_>]$, for any wff A and $w \in W$, $V(\Box A, w) = 1$ just in case $V(A, w') = 1$ for all w' such that wRw'. Otherwise, $V(\Box A, w) = 0$. But the latter is precisely the condition $[V_\Box]$ for a T-model. On the other hand, we may introduce '>' into T by the definition:

$$\text{`}(AB > C)\text{'} =_{df} \text{`}(A \wedge \Box((A \wedge B) \supset C))\text{'}.$$

Notice that by $[V_\wedge]$, $[V_\supset]$, and $[V_\Box]$, $V((AB > C), w) = 1$ iff $V(A, w) = 1$ and for all w' such that wRw', if $V(A, w') = V(B, w') = 1$, then $V(C, w') = 1$. Otherwise $V((AB > C), w) = 0$. This, of course, is the condition $[V_>]$. Since the conditions on W and R are the same for S- and T-models, we may conclude that S- and T-models coincide. Thus, by the respective completeness of S and T (see especially the Corollary above), S and T are equivalent.

That R2 is nothing more nor less than the rule of necessitation explains and motivates R2. Ax>3 turns out to be the straightforward T-theorem, $((A \wedge \Diamond(A \wedge B \wedge \sim C) \wedge (E_1 \wedge \Box((E_1 \wedge F_1) \supset G_1)) \wedge \ldots \wedge (E_n \wedge \Box((E_n \wedge F_n) \supset G_n))) \supset \Diamond(A \wedge B \wedge \sim C \wedge ((E_1 \wedge F_1) \supset G_1) \wedge \ldots \wedge ((E_n \wedge F_n) \supset G_n)))$. We have already seen how Ax>3 pays its way in establishing the Completeness Theorem.

5 ***Comparison with other systems*** Recent discussion of the subjunctive or counterfactual conditional divides into two phases. The first, the so-called "consequence theory",[6] appears in Chisholm [4], Goodman [8], Mackie [14], and Rescher [20]. The second, the possible worlds with similarity view, was developed independently by Stalnaker [21] and Thomason [22] on the one hand, and by Lewis [10] [11] on the other. This approach is also employed by Åqvist [1], Nute [16] [17], and Pollock [19].

According to Chisholm and Goodman, 'If A were the case, then C would be the case' is true if and only if there is a statement or set of statements S such that

$$[\text{entail–Chisholm}]^7 \quad \frac{A \wedge S}{\overline{}\; C.} \quad [\text{lead by law–Goodman}]$$

Both agree that S must be true and must satisfy special conditions. For Goodman, what "lead by law" means must also be qualified. Neither Chisholm nor Goodman can specify what these conditions are, and so their analyses remain incomplete. The central difficulty is that accidental generalizations must be disallowed as members or conjuncts of S for Chisholm, or as candidates for laws to lead to C for Goodman. However, distinguishing accidental generalizations from genuine laws evidently appeals to the notion of a counterfactual, and so the procedure appears circular.

Mackie's account [14] may be represented by the same schema. However, for Mackie a counterfactual is not true if such an argument exists, but rather is a "condensed", "telescoped" presentation of the argument. Although we acknowledge the existence of additional premises, we may not be able to say exactly what they are. Counterfactuals in general then are not true or false but are sustained by the set of implicit premises.

Lewis [11] does not reject the consequence theory as presented by Chisholm and Goodman (although he does criticize Mackie's version), but rather tries to accommodate its insights and solve the outstanding problems.[8] His attitude is conciliatory, not antagonistic. Lewis' possible world system for the semantical analysis of counterfactuals is well-known. The reader may consult [11], especially Chapter 1, for an exposition of his views. His introduction of a similarity relation holding among possible worlds has occasioned serious objections, however. If a world b is more similar to a than is c, what intuitively should be true of a, b, and c? It might seem that more propositions are jointly true of a and b than they are of a and c. And by this we mean that the set of propositions true of a and c is a proper subset of the propositions jointly true of a and b. Our intuitions here are of the same stripe as Gabbay's:

> Remember that we can tell the difference between worlds only through sentences of our language, so we have no choice (particularly in the case of the propositional calculus!) but to equate a world with its complete theory (the set of all sentences true at that world). Well, what does it mean, then, that two theories are similar? Is it that they have, more or less, the same true sentences? ([7], pp. 99-100)

Fine [6] and Bennett [2] have this intuition in mind when they object that "If Oswald had not shot Kennedy, someone else would have" is true on Lewis' view. For it seems that the world where a second assassin kills Kennedy is closer to our world than one where Kennedy survives.[9]

Lewis takes this objection seriously, and in [12] tries to meet it. His program is to argue that there may be many different concepts of similarity among possible worlds. The appropriate relation for evaluating counterfactuals must be chosen from these, and there are priorities which must be met when choosing this relation. If a world violates a condition higher in the priority, then no matter how intuitively similar it may be to a given world, it will not count as being similar to that world in the appropriate respect as some other world intuitively less similar which violates only conditions of lower priority. Merely having more propositions jointly true of a and b rather than a and c is not sufficient for saying that a is more similar to b than to c. If there are major variations in law between a and b or the spatio-temporal region of perfect similarity between a and b is smaller than that between a and c (as in the Kennedy example), a must be counted more similar to c on Lewis' view. This is so, even if a and b share more particular facts.

However in setting up his system of priorities, we feel Lewis forsakes overall similarity between worlds for similarity in certain respects. His theory then does not differ from the consequence view as radically as one might at first suspect. In fact, a view which adopted a certain ranking of priorities and the consequence view might even converge.[10]

We feel that our view combines the best of both the consequence and possible worlds *cum* similarity approaches. Clearly the critical discussion in [2], [6], and [24], and the clarification of Lewis' theory in [12] indicate that the consequentialist intuition is correct in this respect: In asserting a counterfactual conditional, we appeal to certain features of the way the world is. This appeal is selective. We do not appeal to all features of the world and will clearly consider many irrelevant. On the other hand, we agree with Stalnaker and Lewis that the subjunctive conditional is a connective, not a metalinguistic relation.

These intuitions of both parties are combined and preserved in the ternary modal subjunctive conditional $RP > Q$. The R incorporates selection of features of the world. Notice, however, that unlike the consequence view, given P and Q, we are not seeking some fixed R such that $P > Q$ will be true if R, P, and Q satisfy certain conditions. This eliminates the problem of specifying the conditions R is to satisfy which plagued the consequence view. On the other hand, making the similarity condition a syntactic rather than semantic feature of the system allows use of a possible-worlds semantics without a similarity apparatus. The semantic valuation function picks out the set of worlds in which R is true. These then are the worlds sufficiently similar to w to be considered in evaluating the subjunctive conditional. The syntactic component R enables the valuation function to be a similarity function as well.

Possible-worlds *cum* similarity theorists have presented various axiomatizations of the subjunctive conditional. These systems differ in certain significant respects which we shall discuss below. It is interesting to compare these axiomatizations with our view. Nute [17] calls the following system C:

(A1) *Any tautology of the propositional calculus is an axiom.*
(A2) $\Box(A \supset B) \supset (\Box A \supset \Box B)$
(A3) $\Box(A \supset B) \supset (A > B)$
(A4) $\Diamond A \supset [(A > B) \supset \sim(A > \sim B)]$
(A5*) $[A > (B \supset C)] \supset [(A > B) \supset (A > C)]$
(A6) $(A > B) \supset (A \supset B)$
(A7) $(A \gtrless B) \supset [(A > C) \supset (B > C)]$.

The rules of inference are modus ponens for $\supset$ and necessitation in the form $\vdash A$, hence $\vdash \Box A$. (Here '$\Box A$' is defined as '$\sim A > A$' and '$A \gtrless B$' as '$(A > B) \wedge (B > A)$'.) C is the system $C2$ of Stalnaker [22] with the axiom schema

(A5) $[A > (B \vee C)] \supset [(A > B) \vee (A > C)]$

replaced by (A5*).

The reader may verify that A1-A7 with A5* replacing A5 are verified on Lewis' semantics. Also, A5* is valid on Stalnaker's semantics. Hence, C captures a core of intuitions about the subjunctive conditional. Our theory preserves these intuitions, which we take as a mark in its favor. By the definition of $\Box$ and A6, it is trivial to show that $\Box A \supset A$ is a C theorem, establishing that C incorporates a T theory of necessity. Hence by the results of Section 4, S is captured in C. On the other hand, A1 and A2 hold in S.

(A3*) $(C \wedge \Box(A \supset B)) \supset (CA > B)$

is clearly valid. However $(C \wedge \Box(A \supset B))$ says something more than A strictly implies B. The intuition behind A3 is that strict implication is a sufficient condition for the conditional. We may preserve this intuition by noting that in both C and S we have $\vdash \Box(A \supset B) \equiv (t \wedge \Box(A \supset B))$.

(A3)** $\Box(A \supset B) \supset (tA > B)$

may surely be read that if A strictly implies B, then if A were the case, B would be the case, which is what A3 symbolizes.

(A4*) $\Diamond(C \wedge A) \supset [(CA > B) \supset \sim(CA > \sim B)]$
(A5)** $[DA > (B \supset C)] \supset [(DA > B) \supset (DA > C)]$
(A6*) $(CA > B) \supset (A \supset B)$

hold in S. Where '$(DA \gtrless B)$' stands for '$((DA > B) \wedge (DB > A))$',

(A7*) $(DA \gtrless B) \supset [(DA > C) \supset (DB > C)]$

is clearly S-valid.

Three schemata are frequently mentioned in evaluating the systems presented by Stalnaker [21], Lewis [11], and Nute [17]. The first two are intuitively objectionable and counterexamples may be easily constructed. The third has received much discussion recently in the literature. These schemata are

(1) $(A > B) \vee (A > \sim B)$
(2) $(A \wedge B) \supset (A > B)$
(3) $((A \vee B) > C) \supset (A > C)$.

Stalnaker's *C2* contains (1). It is an easy consequence of A5 in the presence of the rule of necessitation, A1, and A3.

(1*) $(CA > B) \vee (CA > \sim B)$

is not semantically valid in our system. Not only may C be false, but in all accessible worlds where both C and A are true, it need not be the case that either B is true or $\sim B$ is true for an arbitrary B. It is easy to see that the analogue of A5,

(A5+) $(DA > (B \vee C)) \supset [(DA > B) \vee (DA > C)]$

also fails. This is of special philosophical interest, since in *C2*, A4 and A5 together imply

(α) $\Diamond A \supset [\sim(A > B) \equiv (A > \sim B)]$.

According to Stalnaker, (α) explains the fact "that the normal way to contradict a counterfactual is to contradict the consequent, keeping the same antecedent" ([21], p. 107). It is easy to see that

(α+) $\Diamond(C \wedge A) \supset [\sim(CA > B) \equiv (CA > \sim B)]$

fails, even assuming C to be true.

Does this indicate a fact about the subjunctive conditional which our analysis cannot account for? We think not. If C and A can both be true together, then $(CA > B)$ and $(CA > \sim B)$ are genuine contraries. Although an effective way of denying one would be to assert the other, they are not contradictories.[11] As it is no fault of a theory of categorical propositions that E and A categoricals should fail to be contradictories, so it is not a fault of our theory that $(CA > B)$ and $(CA > \sim B)$ should fail to be contradictories.

Lewis' system omits (1) but includes (2). Parallel counterexamples appear in [19], [2] (pp. 387-388), and [6] (p. 453). They turn on the intuition that if $A > B$ is true, there has to be some connection between A and B, and this connection must be stronger than both A and B being true. Our system clearly rejects the appropriate analogue of (2):

(2*) $(C \wedge A \wedge B) \supset (CA > B)$.

Clearly C, A, and B may all be true at a world w, but there may be an accessible world where C and A are true but B is false.

Interestingly, Lewis himself is not as sure of (2) as he is of certain other inferences connected with true antecedents, and in the latter cases, our analysis agrees with his. According to Lewis

(a) $(A \wedge \sim B) \supset \sim(A > B)$
(b) $(A > B) \supset (A \supset B)$
(c) $((A > B) \wedge A) \supset B$

are all valid. It is easy to see that their analogues

(a*) $(A \wedge {\sim}B) \supset {\sim}(CA > B)$
(b*) $(CA > B) \supset (A \supset B)$
(c*) $((CA > B) \wedge A) \supset B$

are all valid given our truth condition for $>$.

(3) is genuinely problematic, especially because of its consequences, and intuitions divide on whether it is acceptable. (3) is A8 in Nute [17]. That (3) is rejected by the semantics of both Stalnaker and Lewis can be easily verified. However, as Nute points out in [17], p. 775, when we keep in mind that the disjunction expressed by $\vee$ is *inclusive,* not exclusive disjunction, (3) is a plausible principle for counterfactuals. An analogue of (3),

(3*) $(D(A \vee B) > C) \supset (DA > C)$

is valid according to our semantics.[12]

Professor Dunn has privately offered the following putative counterexample to (3). Nute discusses it in [18], p. 324.

> If Jones were to run for the House or the Senate, then
> he would run for the Senate
> ---
> Therefore, if Jones were to run for the House, then
> he would run for the Senate.

Dunn's example is interesting, but we feel that there is an equivocation between premise and conclusion. This equivocation allows various ways to parse the premise and conclusion, on both the binary and ternary accounts, to show that there is not a genuine counterexample here. Clearly the premise suggests that Jones has a choice between running for the House or Senate, while the conclusion lacks any such suggestions. Letting '*C*' symbolize 'Jones has a choice', '*A*' symbolize 'Jones runs for the House', and '*B*' symbolize 'Jones runs for the Senate', using the binary connective the argument might be rendered:

$$\frac{[C \wedge (A \vee B)] > B}{A > B}.$$

But here the premise is clearly not of the form of the antecedent of (3).

On the other hand, using our ternary connective, we may think of *C* as stating a condition true of the actual world which is tacitly assumed along with certain other facts including that when Jones is given a choice between a more and a less prominent position, he chooses the more prominent. Letting '*D*' symbolize the conjunction of these assumptions, the premise should be symbolized as

$$D(A \vee B) > B.$$

But the conclusion should be symbolized as

$$EA > B.$$

The conclusion does not follow from the premise, but this is no mark against (3*). Notice that on our view,

$$\frac{D(A \vee B) > B}{DA > B}$$

is valid. But this is not counterintuitive. If Jones is given a choice between House and Senate and he would choose the Senate, then if he were to run for the House, he would also run for the Senate. Jones would just be prudent, trying to maximize his chances of getting elected to some public office. In this case, where the assumptions are kept the same from premise to conclusion, the counterexample loses its force.

In [6], p. 453, Fine agrees that the principle seems plausible, but points out the following difficulty: Clearly $A \equiv [(A \wedge D) \vee (A \wedge \sim D)]$, and so

(i) $\quad (A > C) \supset ([(A \wedge D) \vee (A \wedge \sim D)] > C).$

Hence, by (3)

(ii) $\quad (A > C) \supset [(A \wedge D) > C].$

But it seems easy to find counterexamples to (ii). Lewis supplies plenty in [12], p. 10. 'If John had come, it would have been a good party' might be true; but 'if John and Mary had both come, it would have been a good party' might be false, due to the intense animosity between John and Mary. Fine suggests certain ways out of the problem.[13] But our view handles it easily. When one asserts 'If John had come, it would have been a good party', he assumes that no one like Mary came or was going to come. So the statement should be symbolized as '$(DA > C)$', where 'D' includes this information. In asserting 'If John and Mary had both come, it would not have been a good party', the set of implicit assumptions is shifted. Hence the statement should be symbolized by '$(E(A \wedge B) > \sim C)$'. '$(D(A \wedge B) > C)$' might very well be vacuously true. However '$(D(A \wedge B) > C)$' and '$(E(A \wedge B) > \sim C)$' do not contradict each other. But from (3*) we may infer

(δ) $\quad (DA > C) \supset (D(A \wedge B) > C),$

not

(δ*) $\quad (DA > C) \supset (E(A \wedge B) > C),$

and (δ*) is clearly the unacceptable schema.[14]

One further remark is in order. Nute points out [17] that A7 seems to involve us in a paradox. However, it is one which our theory avoids. Since A and $A \vee (A \wedge \sim B)$ are logically equivalent, by A3 and A7, $A > B$ entails $(A \vee (A \wedge \sim B)) > B$. But this seems to have counterintuitive substitution instances. "'If this match were struck, then it would light' may be true, but 'If this match were struck or it were struck and failed to light, then it would light' is prima facie implausible. No normal English speaker who made the first assertion would, after due consideration, make the second" ([17], pp. 776-777).

But what does this amount to on our theory? Certainly $tA \gtrsim (A \vee (A \wedge \sim B))$ holds, and so $tA > B$ entails $t(A \vee (A \wedge \sim B)) > B$. But where A and B symbolize the two statements in the above counterexample, would one contend that $tA > B$ is true? Matches do not always light when struck, but only when they satisfy some further conditions. Hence no normal speaker of English would assert $tA > B$, were he sensitive to the tacit assumptions he was making. On the other hand, should C express such assumptions and C be true, $CA > B$ would entail $C(A \vee (A \wedge \sim B)) > B$. But here again, we feel that no paradox is involved, for there are no worlds where C and $(A \wedge \sim B)$ are true.

Of the analyses of the conditional presented by Stalnaker, Lewis, and Nute, our theory agrees most clearly with that of Nute. This is not surprising, since for Nute "$A > B$ is true in world i just when B is true in all those worlds in which A is true that are *enough like* i for consideration" ([17], p. 774). It is in explicating the notion of one world's being enough like another for consideration that our theories diverge. For Nute, similarity is fixed semantically by a class-selection function. For each world i and wff A, $f(A, i)$ picks out a set of worlds subject to certain conditions. $(A > B)$ is true if B holds at all worlds in $f(A, i)$. (A holds in all such worlds.) For us, a class of worlds is also picked out in evaluating the conditional, $(CA > B)$, namely those accessible worlds in which C is true. The difference between the two approaches is this: For Nute, given a class of conditionals all of the form $(A > B)$ and a world i, the class of worlds at which all such conditionals are to be evaluated is fixed. For us, the classes will vary, since we shall be evaluating expressions of the form $(CA > B)$.

The advantage of our approach lies in this flexibility. Surely for a world j to be enough like i for consideration, certain conditions holding in i must also hold in j. Implicit in Nute's truth condition for counterfactuals is the recognition that in asserting counterfactuals, we make implicit assumptions. However, is it true that any time two conditionals with the same antecedent are asserted, the same implicit assumptions are made? Nute's semantics model this case, but clearly it is a special case. For Nute, if two parties are disputing, one saying that if A were the case, B would be, the other saying that if A were the case, $\sim B$ would be, one of the two must be wrong. On our view, both could be right, if we showed that the implicit assumptions of the two speakers differed widely enough. Hence, the variability of the similarity classes lends a distinct advantage to our approach.

In this connection, we should mention that Gabbay [7] argues that not just the antecedent A, but both A and the consequent B, besides certain features of the world, must be taken into account in determining a class of worlds. Given a world t, A, B, and t determine a set of sentences $\Delta(A, B, t)$. '$(A > B)$' is true at t just in case '$(A \supset B)$' is true at all worlds in which $\Delta(A, B, t)$ is true. Here again, unlike our approach, the set of worlds is fixed.[15]

6 ***Transitivity and contraposition*** As both Stalnaker [21] and Lewis [11] note, transitivity and contraposition for the subjunctive conditional fail on some occasions. They are not universally valid modes of inference. Our theory that subjunctive conditionals involve a suppressed assumption explains why. Consider the following argument: (i) If Churchill had been in Hitler's bunker at the end of the war, then Hitler would have had Churchill put to death. (ii) If Churchill had been Hitler's butler at the end of the war, then Churchill would have been in Hitler's bunker at the end of the war. Therefore, (iii) if Churchill had been Hitler's butler at the end of the war, then Hitler would have had Churchill put to death.

Concerning (i) we probably have these points of similarity to the actual world in mind: Churchill is the loyal British leader. Hitler is the loyal German leader. Britain is at war with Germany. Concerning (ii), however, we probably have these points of similarity to the actual world in mind: Hitler is the loyal German leader. Germany is at war with the Allies. Churchill is loyal to those he serves. Concerning (iii), it is ever so easy to make up points of similarity with the actual world that make it false.

Or consider the following argument: (i) If Churchill were not to have given up smoking, then Churchill would not have been Hitler's butler at the end of the war. Therefore, (ii) if Churchill had been Hitler's butler at the end of the war, Churchill would have given up smoking.

(i) may be true while (ii) is false. Concerning (i), we perhaps have the following in mind: Hitler did not tolerate smoking in his subordinates; Churchill was loyal to those he served. Yet with respect to (ii), we may have the following in mind: Churchill's loyalties do not extend to madmen who have no appreciation of a good cigar.

Lewis [11] points out that under special circumstances, transitivity is valid. What are the conditions under which transitivity and contraposition may be satisfied on our theory?

Any wff of the form $(((XA > B) \wedge (YB > C)) \supset (ZA > C))$ will be satisfied when $((X \wedge Y) \supset Z)$ is satisfied together with any one of the following: (a) $\Box((Z \wedge A) \supset C)$, (b) $\Box((Z \wedge A) \supset (Y \wedge B))$, or (c) $(\Box((Z \wedge A) \supset X) \wedge \Box((Z \wedge A \wedge X \wedge B) \supset Y))$. For example, $(((XA > B) \wedge (XB > C)) \supset (XA > C))$ represents not only a satisfiable, but a valid wff of S.

Any wff of the form $((XA > B) \supset (Y{\sim}B > {\sim}A))$ will be satisfied when $(X \supset Y)$ is satisfied together with one of the following: (a) $\Box((Y \wedge {\sim}B) \supset {\sim}A)$, or (b) $\Box((Y \wedge {\sim}B) \supset X)$. Here $((XA > B) \supset (X{\sim}B > {\sim}A))$ is a valid wff of S.

7 ***Conclusion*** In this paper we have offered an analysis of the subjunctive conditional that treats these conditionals as having a suppressed component. It is our view that when a sentence of the form 'If it were the case that P, then it would be the case that Q' is uttered, its speaker has in mind certain points of similarity that a possible world must have to the actual world to count in the evaluation of what he is saying. This implicit component we have made explicit in our analysis.

Our construing the subjunctive conditional as involving a suppressed component, besides agreeing with the consequence view, also is in line with the intuitions of Mill, Ramsey, Myhill, and Stevenson concerning the implication or conditional connective.[16] We may take this as further data for our view. However, the reader may still have some further objections or difficulties. First, Lewis has presented the subjunctive conditional as a primitive, even *sui generis* connective. Does the fact that our ternary subjunctive conditional '$(AB > C)$' is equivalent to the modal/categorical '$(A \wedge \Box((A \wedge B) \supset C))$' show that the former approach is somehow wrongheaded? Lewis was not wrongheaded in rejecting adamantly any attempt to reduce a counterfactual '$(A > B)$' to a strict conditional '$\Box(A \supset B)$', since on that account if '$(A > C)$' is true, so is '$((A \wedge B) > C)$'. We agree with Lewis that the counterfactual operator is not a strict conditional function of two propositional arguments. The presence of 'A' in '$(AB > C)$' signifies agreement here. Since the 'A' may vary, we also may agree that the counterfactual is a variably strict conditional. Where we disagree and where, as we have argued above, we feel Lewis may be wrongheaded, is in how this element of variability is properly treated. Lewis treats variability through overall similarity. We favor talking about similarity in selected respects, those that are picked out by 'A'.

Secondly, one might object that a speaker may not view his assertion as false if he discovered that some tacit assumption he had in mind was not true of the actual world. This might happen especially if he asserted a counterfactual with an impossible antecedent. In [11], Lewis holds that our intuitions, although not determining the matter decisively, suggest that such counterfactuals are vacuously true. Is "If it were both to rain and not rain, then George Wallace would be the next President of the United States" vacuously true? Since on our analysis, it is symbolized by $(RP > Q)$, where R may be false, it is not.

If the objection is accepted, it is an easy matter to adjust our manner of evaluating '$(RP > Q)$'. One way is simply to omit the first condition of $[V_>]$, i.e., that $V(R, w) = 1$. Thus '$(RP > Q)$' becomes equivalent to '$\Box((R \wedge P) \supset Q)$' in T rather than to '$(R \wedge \Box((R \wedge P) \supset Q))$'. Syntactically, we need merely delete Ax>1. We leave to the interested reader the verification that such a system is sound and complete, along with the assessment of its philosophical merits vis-a-vis the system S presented here and the other systems we have discussed.

Another way, suggested by Professor Jennings, is to change $[V_>]$ to read: for any wffs A, B, and C, and for any $w \in W$, $V((AB > C), w) = 1$ if either: (a) $V(A, w) = 1$ and for every $w' \in W$ such that wRw', if $V(A, w') = V(B, w') = 1$, then $V(C, w') = 1$; or (b) $V(A, w) = 0$ and for every $w' \in W$ such that wRw', if $V(A, w') = 0$ and $V(B, w') = 1$, then $V(C, w') = 1$; otherwise $V((AB > C), w) = 0$.[17] This version of $[V_>]$ does not require that R be true in a world for the subjunctive conditional $(RP > Q)$ to be true in it. Yet the truth or falsehood of R in a world is relevant to the truth of $(RP > Q)$ in it, in that the class of accessible worlds that is relevant to the evaluation of $(RP > Q)$ in, say, the actual world is determined by their similarity to the actual world: either R is true in the actual world and wherever R is true in an accessible world so is $(P \supset Q)$, or R is false in the actual world and wherever R is false in an accessible world $(P \supset Q)$ is true. And as an extra bonus feature, Jennings' suggested interpretation validates $(RP > P)$, read "If it were the case that P, then it would be the case that P". Again, we leave it to the interested reader to explore the details and assess the merit of Professor Jennings' proposal.

NOTES

1. Professor Kit Fine has especially raised this objection with us.
2. Professor W. K. Warmbrōd has communicated a nice example to us. "Most of us have heard enough about the effects of nuclear explosions to know that the counterfactual... is probably true, i.e., 'If full-scale nuclear war were to occur, then the survivors would procreate only monsters'. Yet few of us could actually state the biophysical laws which this counterfactual seems to assume."

3. Professor John Barker has called our attention to a number of these questions. Professor Robert L. Martin has also questioned the variability of the conditional on our account.

4. In a letter, Professor W. K. Warmbrōd has pointed out to us how in ordinary language we may avoid using indexical expressions by replacing them with proper names or definite descriptions. This is an ordinary language device for making these suppressed elements explicit. Professor Warmbrōd points out that our ternary subjunctive conditional extends English in an analogous way, by giving a device for perspicuously expressing the tacit assumptions one makes in stating a subjunctive conditional. This adds to the intuitive appeal of our ternary analysis.

5. See p. 650, for a statement of how this truth condition differs from Gabbay's truth condition for the subjunctive conditional connective in [7].

6. We are here borrowing a phrase of Bennett [2]. Lewis [11] calls this the metalinguistic theory.

7. There are object language/metalanguage problems in assessing just how Chisholm wants tentatively to construe the counterfactual conditional. Introducing notation Chisholm does not use, at one point he suggests that '$(A > B)$' be construed as '$(\exists p)(p \wedge \Box((p \wedge A) \supset B))$' ([4], 298-299). However he also suggests that the conditional '$(x)(y)$ (if x were ϕ and y were ψ, then y would be χ)' "may be rendered as: 'There is a true statement p such that: p and "x is ϕ and y is ψ" entail "y is χ"' ", where a number of conditions must be imposed on p ([4], pp. 300-301). But in the latter, "entails" is not construed as an object language connective, as in the former (necessary material implication), but rather as a metalinguistic relation between statements.

8. Whether Lewis is successful in accommodating the consequence view is another matter. For a discussion of his construal of cotenability, see Bennett [2], Section 6. Fine also maintains that no metalinguistic theorist would accept Lewis' account of cotenability, and that the two approaches differ significantly (see [6], pp. 451-452). Carr ([3], p. 404) also mentions this point.

9. This notion of similarity between worlds also motivates the discussions of Tichý [24] and Pollock [19]. For the latter, see the discussion of minimal change in [19], p. 473.

10. Compare Bennett [2], pp. 397-402.

11. This fact was recognized by Goodman. See [8], p. 115, n. 2, and p. 120, n. 8.

12. In [5], pp. 342-343, Creary and Hill complain that failure to validate (3) makes Lewis' semantics assign true to some conditionals with disjunctive antecedents which are clearly false.

13. See [6], pp. 453-454.

14. This problem which Fine has mentioned has caused philosophers to be suspicious of (3). Nute claims that adding (3) to C reduces the subjunctive conditional to strict implication, allowing these apparently paradoxical inferences ([17], p. 776). Lewis rejects (3) for this reason. In [17], Nute suggests blocking the inference by not allowing substitutivity of provable equivalents in the antecedent of a counterfactual. Fine considers this possibility too. But alas, intuitions do not agree here. Creary and Hill [5] regard (3) as so plausible that this price must be paid. Loewer [13] regards it as too high a price.

15. As with our project, Gabbay gives a semantics with possible worlds but no similarity relation or selection function. However, for Gabbay, the subjunctive conditional is still a binary, not a ternary operator.

16. See [4], p. 298, [15], and [23].

17. A complete axiomatization of a system using Jennings' version of $[V_>]$ can be gained by adding to the tautologies of the propositional calculus (expressed in terms of $\sim$ and $\supset$, the other connectives being defined as usual), the following two axiom schemes:

AxJ1 $(X \wedge (AB > C)) \supset ((X \wedge B) \supset C)$, *where X is A or X is* $\sim A$

AxJ2 $(X \wedge \sim(AB > C) \wedge Y_1 \wedge (E_1F_1 > G_1) \wedge \ldots \wedge Y_n \wedge (E_nF_n > G_n)) \supset \sim(tt > \sim(X \wedge B \wedge \sim C \wedge ((Y_1 \wedge F_1) \supset G_1) \wedge \ldots \wedge ((Y_n \wedge F_n) \supset G_n)))$, *where X is A or X is* $\sim A$, Y_1 *is* E_1 *or* Y_1 *is* $\sim E_1, \ldots, Y_n$ *is* E_n *or* Y_n *is* $\sim E_n$.

The rules of inference remain as in *S*.

REFERENCES

[1] Åqvist, L., "Modal logic with subjunctive conditionals and dispositional predicates," *Journal of Philosophical Logic*, vol. 2 (1973), pp. 1-76.

[2] Bennett, J., "Counterfactuals and possible worlds," *Canadian Journal of Philosophy*, vol. 4 (1974), pp. 381-402.

[3] Carr, B., "Review of *Counterfactuals*," *British Journal for the Philosophy of Science*, vol. 27 (1976), pp. 403-405.

[4] Chisholm, R. M., "The contrary-to-fact conditional," *Mind*, vol. 55 (1946), pp. 289-307.

[5] Creary, L. G., and C. S. Hill, "Review of Counterfactuals," *Philosophy of Science*, vol. 42 (1975), pp. 342-344.

[6] Fine, K., "Critical notice of *Counterfactuals*," *Mind*, vol. 84 (1975), pp. 451-458.

[7] Gabbay, D. M., "A general theory of the conditional in terms of a ternary operator," *Theoria*, vol. 38 (1972), pp. 97-104.

[8] Goodman, N., "The problem of counterfactual conditionals," *The Journal of Philosophy*, vol. 44 (1947), pp. 113-128.

[9] Hughes, G. E., and M. J. Cresswell, *An Introduction to Modal Logic*, Methuen, London, 1968.

[10] Lewis, D., "Completeness and decidability of three logics of counterfactual conditionals," *Theoria*, vol. 37 (1971), pp. 74-85.

[11] Lewis, D., *Counterfactuals*, Basil Blackwell, Oxford, 1973.

[12] Lewis, D., "Counterfactual dependence and time's arrow," *Noûs*, vol. 8 (1979), pp. 455-476.

[13] Loewer, B., "Counterfactuals with disjunctive antecedents," *The Journal of Philosophy*, vol. 73 (1976), pp. 531-537.

[14] Mackie, J. L., "Counterfactuals and causal laws," in *Analytical Philosophy*, ed., R. J. Butler. Basil Blackwell, Oxford, 1962.

[15] Myhill, J., "On the interpretation of the sign '$\supset$'," *The Journal of Symbolic Logic*, vol. 18 (1953), pp. 60-62.

[16] Nute, D., "Counterfactuals," *Notre Dame Journal of Formal Logic*, vol. 16 (1975), pp. 476-482.

[17] Nute, D., "Counterfactuals and the similarity of worlds," *The Journal of Philosophy*, vol. 72 (1975), pp. 773-778.

[18] Nute, D., "Simplification and substitution of counterfactual antecedents," *Philosophia*, vol. 7 (1978), pp. 317-326.

[19] Pollock, J., "The 'Possible Worlds' analysis of counterfactuals," *Philosophical Studies*, vol. 29 (1976), pp. 469-476.

[20] Rescher, N., *Hypothetical Reasoning*, North-Holland, Amsterdam, 1964.

[21] Stalnaker, R. C., "A theory of conditionals," pp. 98-112 in *Studies in Logical Theory*, ed., Nicholas Rescher, Basil Blackwell, Oxford, 1968.

[22] Stalnaker, R. C., and R. Thomason, "A semantic analysis of conditional logic," *Theoria*, vol. 36 (1970), pp. 23-42.

[23] Stevenson, C. L., "If-iculties," pp. 279-309 in *Logic and Art: Essays in Honor of Nelson Goodman*, ed., R. Rudner and I. Scheffler, Bobbs-Merrill, Indianapolis, 1972.

[24] Tichý, Pavel, "A counterexample to the Stalnaker-Lewis analysis of counterfactuals," *Philosophical Studies*, vol. 29 (1976), pp. 271-273.

Charles B. Daniels
20/3987 Gordon Head Road
Victoria, B.C., Canada

James B. Freeman
Hunter College of the City
University of New York

and

Philosophy Department
University of Victoria
Victoria, British Columbia
Canada V8W 2Y2