

## Counterfactuals without Possible Worlds?

### A Difficulty for Fine's Exact Semantics for Counterfactuals

#### [Penultimate Draft]

Thanks in large part to the work of Robert Stalnaker and David Lewis, philosophers standardly use possible worlds semantics for counterfactuals.<sup>1</sup> Roughly, the counterfactual from A to C ( $A > C$ ) is true at  $w$  if and only if the closest worlds to  $w$  in which A is true are worlds in which C is true. In a pair of recent papers, Kit Fine argues that possible worlds semantics for counterfactuals faces grave difficulties.<sup>2</sup> For according to possible worlds semantics, substitution of logically equivalent statements into the antecedent or consequent position of a counterfactual statement preserves truth-value,<sup>3</sup> more precisely, if A is logically equivalent to A', then from  $A > C$  you may infer  $A' > C$ , and if C is logically equivalent to C', then from  $A > C$  you may infer  $A > C'$ . Call this principle "Substitution." As is well known, Substitution is valid in possible worlds semantics because logically equivalent statements are true in the same worlds. Hence, if the closest A worlds are C worlds, and A' is logically equivalent to A, then the closest A' worlds are C worlds. The same holds, *mutatis mutandis*, for C and C'. Fine argues that Substitution is incompatible with the conjunction of certain other highly intuitive principles of counterfactual reasoning, principles such as

Infinitary Conjunction       $A > C_1, A > C_2, A > C_3, \dots$

$A > C_1 \wedge C_2 \wedge C_3 \wedge \dots$

Disjunction                       $A > C, B > C$

$A \vee B > C$

Transitivity

$$\underline{A > B, A \wedge B > C}$$

$$A > C$$

As an alternative to possible worlds semantics, Fine develops what he calls “exact semantics” and shows that exact semantics avoids the problems of possible worlds semantics while preserving the validity of principles such as Infinitary Conjunction, Disjunction, and Transitivity. However, as I will argue, exact semantics has problems of its own: it yields counter-intuitive results when evaluating counterfactuals with what I call ‘determinable’ antecedents, which can be made true in a variety of ways.

In §1 I present Fine’s exact semantics for counterfactuals, offering clarification where needed. Fine’s discussion is of course much richer than what space allows here: I only present the portion of Fine’s semantics necessary for understanding the difficulty presented in §2. In §3 I consider a response to the difficulty raised in §2. Unfortunately, the response requires giving up either Disjunction or Infinitary Conjunction.

### *1. Exact Semantics for Counterfactuals*

In Fine’s exact semantics,<sup>4</sup> atomic propositions are made true by parts of worlds—something like states, situations, or facts. Here I follow Fine in speaking of states as truthmakers for propositions. The idea is that a state  $s$ , in which  $A$  is the case, makes it true that  $A$ . For example, the state wherein it is snowing makes it true that it is snowing. Moreover, states can be parts of other states. There is a fusion of two or more states just in case the component states are jointly compossible; ‘ $s_1 \sqcup s_2$ ’ denotes the fusion of states  $s_1$  and  $s_2$ . I will call a state ‘composite’ just in case it is a fusion of two or more states.

Both simple and composite states can be truthmakers. I will speak of a truthmaker having parts or being composed of states if and only if the truthmaker is a composite state.

There are several ways to think about the truthmaking relation. First there is a *loose* conception of truthmaking. Loose truthmaking is a merely modal notion:  $s$  makes it true that  $A$  in the loose sense if and only if it is impossible that  $s$  obtain and  $A$  be false. Since it is impossible that necessary truths be false, every state is a truthmaker in the loose sense for every necessary truth, for every state is such that it is impossible for it to obtain while any necessary truth is false. For example, the fact that I am sitting is a loose truthmaker for the proposition that it is snowing or it is not snowing in Toronto.

As its name suggests, Fine's exact semantics does not employ the loose conception of truthmaking. The *exact* conception of truthmaking employs a relevance constraint:  $s$  is an exact truthmaker for  $A$  if and only if it is impossible that  $s$  obtain and  $A$  be false, and  $s$  is wholly relevant to  $A$ . I assume, as Fine does, that we understand what it is for a state to be relevant to a proposition. But what is it for a state to be *wholly* relevant to a proposition? Fine does not say, but a plausible suggestion is that  $s$  is wholly relevant to  $A$  just in case  $s$  is relevant to  $A$  and does not have parts that are irrelevant to  $A$ .<sup>5</sup> Thus, if  $s_1$  is relevant to  $A$ , and  $s_2$  is not, then  $s_1 \sqcup s_2$  is not an exact truthmaker for  $A$ , since it has a part,  $s_2$ , that is not relevant to  $A$ .

Between loose and exact truthmaking, there is inexact truthmaking. A state  $s$  is an inexact truthmaker for  $A$  just in case  $s$  is either an exact truthmaker for  $A$  or an extension of an exact truthmaker for  $A$ , where an extension of an exact truthmaker for  $A$  is composed of multiple states, at least one of which is an exact truthmaker for  $A$ , and at least one of which is not. If  $s_1$  is an exact truthmaker for  $A$  and  $s_2$  is not, then both  $s_1$  and  $s_1 \sqcup s_2$  are inexact truthmakers for  $A$ . It is important that exact truthmakers also be

considered inexact truthmakers, for otherwise  $A > A$  would turn out false on Fine's semantics.<sup>6</sup>

Fine gives the following exact verification conditions for molecular propositions:

- (i)<sup>+</sup> state  $s$  exactly verifies  $\neg A$  iff  $s$  falsifies<sup>7</sup>  $A$ ;
- (i)<sup>-</sup>  $s$  falsifies  $\neg A$  iff  $s$  exactly verifies  $A$ ;
- (ii)<sup>+</sup>  $s$  exactly verifies  $A \wedge B$  iff  $s$  is the fusion  $s_1 \sqcup s_2$  of a state  $s_1$  that exactly verifies  $A$  and a state  $s_2$  that exactly verifies  $B$ ;
- (ii)<sup>-</sup>  $s$  falsifies  $A \wedge B$  iff  $s$  falsifies  $A$  or  $s$  falsifies  $B$  or  $s$  falsifies  $A \vee B$ ;
- (iii)<sup>+</sup>  $s$  exactly verifies  $A \vee B$  iff  $s$  exactly verifies  $A$  or  $s$  exactly verifies  $B$  or  $s$  exactly verifies  $A \wedge B$ ;
- (iii)<sup>-</sup>  $s$  falsifies  $A \vee B$  iff  $s$  is the fusion  $s_1 \sqcup s_2$  of a state  $s_1$  that falsifies  $A$  and a state  $s_2$  that falsifies  $B$ .

Inexact verification can now be defined recursively in terms of exact verification.

To extend the given semantics to a semantics for counterfactuals, Fine introduces the ternary relation ' $t \rightarrow_w u$ ', which he renders as, " $u$  is a possible outcome of *imposing the change*  $t$  on the world  $w$ ." The truth condition for counterfactuals is then stated as follows:

- (TC)  $A > C$  is true at  $w$  if and only if  $u$  inexactly verifies  $C$  whenever  $t$  exactly verifies  $A$  and  $u$  is a possible outcome of  $t$  relative to  $w$ .

The crucial notion of a possible outcome invites three questions. First, what is it for one state to be a possible outcome of another? Second, what does imposing a change on a

world amount to? Third, why must possible outcomes be relative to worlds? In other words, why is the *possible outcome* relation ternary? I will take these questions in turn.<sup>8</sup> Possible outcomes are naturally taken to be what Fine calls “future causal outcomes.” But as Fine notes, not all counterfactuals are true in virtue of a forward-looking causal connection. Examples of such counterfactuals include back-tracking counterfactuals and counterfactuals that track a logical or conceptual relation.<sup>9</sup> An example of the latter is, “if Johnny were to have a child, his brother would be an uncle.”<sup>10</sup> In addition to back-tracking counterfactuals and counterfactuals that track a conceptual relation, defenders of libertarian free will might want to add counterfactuals of freedom such as, “if Curly were offered a bribe, he would freely take it.” Such counterfactuals are not true in virtue of causal connections, or so say libertarians. Without giving the term ‘possible outcome’ a formal treatment, Fine recommends interpreting it broadly enough to accommodate counterfactuals that are true in virtue of such non-causal relations.

How are we to understand imposing a change on a world? Here Fine does not give us much help, but notice that  $w$ , the world being “changed,” is also the world of evaluation of the counterfactual. As is familiar, for  $A > C$  to be true at  $w$ , it is not necessary that either  $A$  or  $C$  be true at  $w$ . Accordingly, neither  $A$  nor  $C$  need have truthmakers at  $w$ . Talk of imposing a change on  $w$ , the world of evaluation, is meant to signal that  $t$  and  $u$ , the truthmakers for  $A$  and  $C$ , need not be states of  $w$ .

Finally, why must possible outcomes be relative to the world of evaluation? Here too Fine offers little clarification, but my suggestion is as follows. When we evaluate a counterfactual  $A > C$  at a world, we hold fixed certain features of that world. “If I had let go, Sue would have fallen” is true in the actual world, given certain causal features of the actual world. Without holding fixed the relevant causal features of the actual world, it is not true that, had I let go, Sue would have fallen: perhaps she would have floated away.

Similarly, when we evaluate whether  $u$  is a possible outcome of  $t$ , we need to hold fixed relevant features of the world of evaluation. Curly's freely taking a bribe is a possible outcome of Curly's being offered a bribe, given relevant features of Curly's character, the circumstances in which he is offered the bribe, and so forth. If those features are not held fixed—if Curly were an upstanding citizen, for instance, or if he knew the feds were watching—he would not freely take the bribe. Relativizing the possible outcome relation to the world of evaluation is perhaps meant to capture the fact that certain features of the world of evaluation are relevant to the truth of the counterfactual being evaluated. These remarks are enough, I hope, to shed some light on the possible outcome relation; however, my objection to Fine's exact semantics does not hinge on the clarifications I have here suggested.

Because molecular propositions can have multiple truthmakers, a question remains about how to apply (TC) to counterfactuals with molecular antecedents and consequents. As (iii)<sup>+</sup> above makes clear, if  $s_1$  exactly verifies  $A$ , and  $s_2$  exactly verifies  $B$ , then  $A \vee B$  is exactly verified by  $s_1$ ,  $s_2$ , and  $s_1 \sqcup s_2$ . What does (TC) have to say about a counterfactual like  $A \vee B > C$ ? Must the right-hand side of (TC) be satisfied for just one of  $A \vee B$ 's exact verifiers? Or must it be satisfied for all of them? According to Fine, all of them:

Our account of the truth-conditions for counterfactuals will be based upon two main ideas. The first is the *Universal Realizability of the Antecedent*. What this means is that a counterfactual  $A > C$  will only be taken to be true when it is true for *any way* in which its antecedent  $A$  might be true. The second idea is the *Universal Verifiability of the Consequent*. What this means is that a counterfactual will be taken to be true, given some way in which its antecedent might be true, only when its consequent is made true under *any outcome* of the way in which its antecedent is true.<sup>11</sup>

Thus (TC) involves universal quantification over states. It will be useful to restate (TC) to make Universal Realizability and Universal Verifiability perspicuous:

(TC')  $A > C$  is true at  $w$  if and only if, for every state  $u$  that is a possible outcome, relative to  $w$ , of any state  $t$  that is an exact verifier of  $A$ ,  $u$  is an inexact verifier for  $C$ .

Note that my (TC') is equivalent to Fine's (TC); it merely makes explicit what is contained in the term 'whenever' in (TC) and makes it easier to see how (TC) is to be applied to counterfactuals with antecedents or consequents that have multiple truthmakers.

Note that Substitution is not generally valid in exact semantics. For instance, from  $A > C$  you may not infer  $(A \wedge B) \vee (A \wedge \neg B) > C$  because it is not the case that every exact verifier of  $(A \wedge B) \vee (A \wedge \neg B)$  is also an exact verifier of  $A$ . Here an example might help.

From

(1a) If Sue were to take her pills, she would get better,

one cannot infer

(1b) If Sue were to take her pills and the cyanide or her pills and not the cyanide, she would get better.

To see why (1b) cannot be inferred from (1a), let  $t$  be the state wherein Sue takes the pills and the cyanide, and let  $u$  be the state wherein Sue dies (as opposed to getting better).  $t$  is an exact verifier for the antecedent of (1b), and  $u$  is a possible outcome of  $t$ , but  $u$  does not inexactly verify the consequent of (1b). Accordingly (1b) is false although (1a) is true.

Universal Realizability of the Antecedent is supported by examples such as,

(2) If Sue were to take her pills or the cyanide, she would get better.

(2) is intuitively false. On Fine's exact semantics, an explanation for (2)'s being false is forthcoming. Let  $t$  be the state wherein Sue takes the cyanide, and let  $u$  be the state wherein Sue dies. Now  $t$  is an exact verifier for the antecedent of (2), and  $u$  is a possible outcome of  $t$ . But  $u$  does not inexactly verify the consequent of (2). Hence, (2) is false. This result might be thought an advantage over possible worlds semantics, which would assign truth to (2), assuming that the worlds in which Sue takes the pills and gets better are closer to the actual world than the worlds in which she takes the cyanide and does not get better.

However, in the next section I argue that Universal Realizability of the Antecedent is too strong. Because Fine's truth condition for counterfactuals is based on Universal Realizability of the Antecedent, it assigns falsehood to a wide range of counterfactuals that are intuitively true.

## *2. A Difficulty for Exact Semantics*

It snowed eight inches yesterday. So it is true that it snowed yesterday. The state wherein it snowed eight inches makes it true that it snowed. If it had snowed more (or less), it would still be true that it snowed. For instance, if it had snowed eight feet instead of eight inches, it would still be true that it snowed. In that case, the state wherein it snowed eight feet would make it true that it snowed. Thus, there are many ways for it to be true that it snowed.



I call an atomic proposition ‘determinable’ if and only if it can be made true by a variety of determinate states. The proposition that it snowed is determinable. So is the proposition that I have more than ten dollars. It can be true that I have more than ten dollars in virtue of my having \$11, and it can be true that I have more than ten dollars in virtue of my having \$11,000. Both states would make it true that I have more than ten dollars.

Determinable propositions present a problem for exact semantics.<sup>12</sup> We may suppose that the following counterfactual is true:

(3a) If it had snowed yesterday, I would have gone skiing.

(3a) is true for many ways in which the antecedent could be true, but not for all of them.

For instance, it is false that

(3b) If it had snowed 100 feet yesterday, I would have gone skiing.

Since 100 feet of snow is one way for the antecedent of (3a) to be true, (3a) is not true for every way in which its antecedent could be true. To see exactly what the difficulty is, let  $t$  be the state wherein it snowed 100 feet yesterday, and let  $u$  be the state wherein I stay home, no doubt disappointed by the cancellation of my ski trip. Now,  $t$  is an exact verifier for the antecedent of (3a), and  $u$  is a possible outcome of  $t$ , but  $u$  does not inexactly verify the consequent of (3a). So according to (TC'), (3a) is false. This is a counter-intuitive result of exact semantics.

Nor is this sort of case isolated. To take another example, we may suppose that (4a) is true:

(4a) If Sue were to take some of these pills, she would get better.

But on the counterfactual supposition that Sue takes 25 of these pills, she would not get better. Thus (4b) is false:

(4b) If Sue were to take 25 of these pills, she would get better.

Since the antecedent of (4b) expresses one way in which the antecedent of (4a) could be true, (4a) must also be false. But intuitively, (4a) is the sort of thing that could be true, notwithstanding what would happen if Sue were to take too many pills.

One more example:

(5a) If Sue were not in the driver's seat, she would have survived.

(5b) If Sue were in between the two colliding cars, she would have survived.

Intuitively, we may suppose that (5a) is true while (5b) is false. But the antecedent of (5b) expresses one way for the antecedents of (5a) to be true. So in exact semantics (5a) turns out false.

The difficulty comes from Universal Realizability of the Antecedent, so an advocate of exact semantics might be tempted simply to jettison this condition. The resulting truth condition for counterfactuals might then be as follows:

(TC'')  $A > C$  is true at  $w$  if and only if there is some state  $t$  that exactly verifies  $A$  and is such that for every state  $u$  that is a possible outcome of  $t$  relative to  $w$ ,  $u$  is an inexact verifier for  $C$ .

Unfortunately, (TC'') will not do. For according to (TC''), (3a) turns out true, but so does

(3c) If it had snowed yesterday, I would not have gone skiing.

(3c) turns out true because it is true for one way in which the antecedent could be true – namely, if it snowed 100 feet. Thus both (3a) and (3c) are true.

(3a) If it had snowed yesterday, I would have gone skiing;

(3c) If it had snowed yesterday, I would not have gone skiing.

The conjunction of (3a) and (3c) might not be a contradiction, but given the plausible principle of Conjunction,

Conjunction       $A > B, A > C$

$A > B \wedge C$

we may infer that if it had snowed yesterday, I would have gone skiing and I would not have gone skiing. In other words, if it had snowed, I would have brought about a contradiction, which is surely implausible. Thus the exact semantics faces difficulties with and without Universal Realizability of the Antecedent.

Possible worlds semantics does not face the same difficulties, for in the possible worlds semantics  $A > C$  is true just in case the closest  $A$ -worlds are  $C$ -worlds. Here the quantifier in the truth condition is existential rather than universal, and there is a constraint, in terms of closeness, on which worlds count for evaluating the counterfactual. Thus (3a) is true because the closest world in which it snowed is a world in which it snowed enough to go skiing but not enough to cause a catastrophe. But (3b) is false because the closest worlds in which it snowed 100 feet are not worlds in which I went skiing. Similar reasoning applies to (4a)-(5b). Possible worlds semantics employs a similarity metric to discriminate between worlds that matter for the truth of the counterfactual and worlds that do not; in exact semantics, by contrast, all ways of exactly verifying the antecedent are on a par. As a result, exact semantics does not yield the correct truth-values for counterfactuals that are true for some but not all ways that their antecedents can be true.

The comparison with possible worlds semantics suggests a solution: combine the approaches. Perhaps we should use similarity between possible worlds to give the truth condition for counterfactuals more power of discrimination between ways in which the antecedents and consequents can be true. How would such an approach go?

### 3. *A Combined Approach*

Here is one suggestion:

(TC\*)  $A > C$  is true at  $w$  iff  $u$  inexactly verifies  $C$  in the world closest to  $w$  in which  $t$  exactly verifies  $A$  and  $u$  is an outcome of  $t$ .<sup>13</sup>

Because (TC\*) employs the notions of exact and inexact verification, it has the (by Fine's lights) desirable result of yielding a semantics according to which Substitution is not generally valid. For instance, given the truth of

$$A > C$$

we cannot infer the truth of

$$(A \wedge B) \vee (A \wedge \neg B) > C$$

because an exact verifier for A is not an exact verifier for  $(A \wedge B) \vee (A \wedge \neg B)$ .

However, (TC\*) as it is cannot be correct, for it relies on the so-called Limit Assumption. Consider the set of worlds in which A is true. Call these worlds the A-worlds. The Limit Assumption says that for every world  $w$  there is a closest A-world. If the Limit Assumption is false, the A-worlds get closer and closer to  $w$  *ad infinitum*, just as the real numbers get closer and closer to 0 *ad infinitum*. Following Fine, we call a set of worlds "stranded" with respect to a base world if and only if the set contains no world that is closest to the base world.<sup>14</sup> The members of a stranded set are stranded worlds.

To see how the Limit Assumption can fail, suppose there is an actual line one inch in length. Call it "Linus." In a nearby world Linus is 1 ½" in length. In a still closer world Linus is 1 ¼" in length. In a still closer world Linus is 1 1/8" in length. Obviously, this can go on *ad infinitum*. So the set of worlds in which Linus is greater than 1" in length is stranded.<sup>15</sup> Using exact semantics, how shall we evaluate a counterfactual against a space of stranded worlds?

Lewis suggests that we take  $A > C$  to be true at  $w$  if and only if there is at least one  $(A \wedge C)$ -world,  $v$ , and there are no  $(A \wedge \neg C)$ -worlds closer than  $v$  to  $w$ .<sup>16</sup> Applying this idea to Fine's exact semantics yields the following proposal.

(TC\*\*)  $A > C$  is true at  $w$  iff there is a world  $v$  in which  $u$  inexactly verifies  $C$  and  $t$  exactly verifies  $A$  and  $u$  is an outcome of  $t$ , and there is no world  $v^*$  closer than  $v$  to  $w$  in which  $(A \wedge \neg C)$  is true.

It is well known that Infinitary Conjunction fails in Lewis's semantics.<sup>17</sup>

Infinitary Conjunction	$A > C_1, A > C_2, A > C_3, \dots$
	$A > C_1 \wedge C_2 \wedge C_3 \wedge \dots$

To see why, return to the case of Linus and the stranded worlds in which Linus is greater than 1'' in length. We can now see that on Lewis's semantics the following counterfactual is true.

(6) If Linus were longer than 1'', Linus would not be longer than 1 1/2''.

(6) is true because there are worlds in which Linus is less than 1 1/2'' in length that are closer to the base world (in which Linus is 1'' in length) than all the worlds in which Linus is greater than 1 1/2'' in length. But by the same reasoning (7) is true.

(7) If Linus were longer than 1'', Linus would not be longer than 1 1/4''.

But this can go on *ad infinitum*. If Infinitary Conjunction were valid, we would be able to infer that if Linus were longer than 1” in length, Linus would not be longer than any length greater than 1”, which is false. The important point for now is that the same argument applies to the exact semantics if (TC\*\*) is the preferred truth condition. Fine argues that the rejection of Infinitary Conjunction is unacceptable.<sup>18</sup> So presumably Fine would not be happy with (TC\*\*).

Fine introduces an alternative to Lewis’s semantics within the possible worlds framework.<sup>19</sup> Instead of taking  $A > C$  to be true at  $w$  if and only if the closest  $(A \wedge C)$ -world is closer to  $w$  than every  $(A \wedge \neg C)$ -world, we take  $A > C$  to be true at  $w$  if and only if *all* of the stranded  $A$ -worlds are  $C$ -worlds. On this alternative semantics, propositions such as (6) and (7) turn out false because not all of the worlds in which Linus is greater than 1” in length are worlds in which Linus is not greater than 1 ½” in length (or 1 ¼” in length, and so on).

The alternative semantics just sketched can be adapted for the work of exact semantics. We will now need two different truth conditions, one for stranded worlds and one for non-stranded worlds.

- (Non-Stranded)  $A > C$  is true at  $w$  iff there is a closest world  $v$  in which  $u$   
inexactly verifies  $C$  and  $t$  exactly verifies  $A$  and  $u$  is an outcome  
of  $t$   
or  
(Stranded) the  $A$ -worlds are stranded with respect to  $w$  and all of them  
contain a state  $u$  that inexactly verifies  $C$  and a state  $t$  that  
exactly verifies  $A$ , where  $u$  is an outcome of  $t$ .

However, as Fine shows, in the alternative semantics the principle of Disjunction is not valid.

Disjunction     $A > C, B > C$

$A \vee B > C$

Consider a set of worlds  $w_1, w_2, w_3, \dots$  getting progressively closer to a base world  $i$ . Suppose that A is true only in  $w_1$  and  $w_2$ , B is true only in the infinitely many worlds  $w_2, w_3, \dots$  leading up to  $i$ , and C is also true only in the infinitely many worlds  $w_2, w_3, \dots$  leading up to  $i$ . Since the A-worlds are not stranded, and the closest A-world ( $w_2$ ) is a C-world,  $A > C$  is true. The B-worlds are stranded, but since all of the B-worlds are C-worlds,  $B > C$  is true. Now the  $(A \vee B)$ -worlds are also stranded, but not all of them are C-worlds ( $w_1$  is an  $(A \vee B)$ -world but not a C-world). So  $A \vee B > C$  is false, and Disjunction is invalid. The important thing to see for present purposes is that the same argument applies to the exact semantics expressed by the truth conditions (Non-Stranded) and (Stranded). As Fine maintains, rejection of Disjunction is a great cost, since Disjunction is intuitive, and we have no independent reason to reject it.<sup>20</sup>

#### 4. Conclusion

I have argued that exact semantics as developed by Fine cannot accommodate counterfactuals with determinable antecedents. The problem is that in exact semantics all ways of making the antecedent true are on a par. In order to remedy this problem, it seems we need recourse to the similarity relation employed in possible worlds semantics. However, this approach results in a dilemma. If we combine the standard possible worlds semantics with exact semantics, Infinitary Conjunction is invalid; if we combine Fine's



alternative approach to possible worlds semantics with the exact semantics, Disjunction is invalid. Both of these results are undesirable. What is more, Fine's argument against possible worlds semantics relies on both Infinitary Conjunction and Disjunction. So if we have to reject one of those principles, we might as well hang on to Substitution and the standard possible worlds semantics.<sup>21</sup>

---

<sup>1</sup> David Lewis, *Counterfactuals* (Cambridge: Harvard, 1973); Robert Stalnaker, "A Theory of Conditionals," in Nicholas Rescher, ed., *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, Number 2 (Oxford, UK: Blackwell, 1968): 98-112.

<sup>2</sup> Kit Fine, "A Difficulty for the Possible Worlds Analysis of Counterfactuals," *Synthese*, CLXXXIX (2012): 29-57; "Counterfactuals without Possible Worlds," *The Journal of Philosophy*, CIX, no. 3, (March, 2012): 221-246.

<sup>3</sup> As Fine notes in "A Difficulty for the Possible Worlds Analysis of Counterfactuals," p. 30, Substitution is valid in possible worlds semantics not only for logically equivalent statements but also for modally equivalent statements—that is, for any statements A and A' (C and C') that are true in all the same worlds. I focus on Substitution as it applies to logically equivalent statements because that version of the principle is weaker, and it is the version of the principle that Fine attacks.

<sup>4</sup> As given in Fine, "Counterfactuals without Possible Worlds," pp. 233-237.

<sup>5</sup> Note that simple states do not have parts, so *a fortiori*, do not have parts that are irrelevant to A. So if s is simple and relevant to A, then s is wholly relevant to A.

<sup>6</sup> See (TC) below.

<sup>7</sup> Fine does not say much about the notion of falsification. For our purposes, we may take falsification to be defined recursively in terms of exact verification.

<sup>8</sup> Thanks to an anonymous reviewer for pressing these questions.

---

<sup>9</sup> For more on back-tracking counterfactuals, see David Lewis, “Counterfactual Dependence and Time’s Arrow,” *Noûs*, XIII, no. 4 (1979): 455-476.

<sup>10</sup> Fine’s example is, “if this peg had been round, it would not have fit the hole,” but without building further conditions into the antecedent and consequent, it is not obvious that this example tracks a conceptual relation, since there is no obvious conceptual relation between being round and not fitting holes.

<sup>11</sup> Fine, “Counterfactuals without Possible Worlds,” p. 236.

<sup>12</sup> Some of the examples that follow were inspired by Crispin Wright’s discussion in “Keeping Track of Nozick,” *Analysis*, XLIII, no. 3 (June, 1983): 134-140.

<sup>13</sup> In what follows I assume that A is entertainable at *w*. I leave that assumption out of the truth conditions presented for the sake of brevity.

<sup>14</sup> Fine, “Counterfactuals without Possible Worlds,” p. 228.

<sup>15</sup> See Lewis, *Counterfactuals*, pp. 19-21.

<sup>16</sup> Lewis, *Counterfactuals*, p. 21.

<sup>17</sup> See John Pollock, *Subjunctive Reasoning*, Philosophical studies series in philosophy, Volume 8, (Dordrecht: D. Reidel, 1979): 18-23; Fine, “Counterfactuals without Possible Worlds,” pp. 225-226.

<sup>18</sup> Fine, “Counterfactuals without Possible Worlds,” p. 226.

<sup>19</sup> Fine, “Counterfactuals without Possible Worlds,” p. 228.

<sup>20</sup> Fine, “Counterfactuals without Possible Worlds,” p. 229.

<sup>21</sup> Thanks to Benj Hellie, Phil Kremer, Adam Murray, Jessica Wilson, and an anonymous reviewer for helpful comments. I want to single out Jessica Wilson for special thanks for her extensive comments and very helpful discussion.