No New Problem for Counterpossibles*

Michael De
University of Utrecht
Department of Philosophy
Utrecht, Netherlands
michael.de@phil.uu.nl

February 8, 2012

1 Introduction

In an article in this journal [Goo04], Jeffrey Goodman proposes what he calls a “new problem for counterpossibles”. A counterpossible is a counterfactual with a metaphysically impossible antecedent. The old problem is that, on standard analyses of counterfactuals due to Lewis [Lew73] and Stalnaker [Sta68], some intuitively false counterpossibles turn out true. The new problem is that, when these standard analyses are extended to deliver non-vacuous truth conditions for counterpossibles, some intuitively true counterpossibles turn out false. I argue that Goodman has not shown there is a genuine problem for counterpossibles.

2 The standard account of counterfactuals

Consider what Goodman calls the ‘Standard Account’ (SA)\(^1\) of counterfactuals:

**Standard Account:** A counterfactual, i.e., a sentence \(R\) of the form ‘If \(A\) were true, then \(B\) would be true’, is true at a world \(\beta\) iff either \(A\) is impossible, or in the closest possible world \(\alpha\) to \(\beta\) where \(A\) is true, \(B\) is also true; otherwise \(R\) is false [at] \(\beta\).\(^2\)

---

* I would like to thank Thomas Müller for helpful discussion on an earlier draft of this paper. This research was funded by the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement nr 263227.

\(^{1}\) All abbreviations and labels are Goodman’s.

\(^{2}\) I have inserted ‘at’ where ‘in’ occurs since I presume this is a typographical error.
SA is a partial gloss on Stalnaker’s semantics.\textsuperscript{3} According to SA, every counterpossible is vacuously true.\textsuperscript{4} But some counterpossibles seem false, such as:

(M) If Saul Kripke were an earthworm, the square root of 9 would be 2.

Why does Goodman think that (M) is false? He says “[t]here is no reason to suppose the square root of 9 would be 2 were Kripke an earthworm[...the antecedent does not seem to be connected to the consequent in the right way to make the antecedent true” [Goo04, p. 43, my emphasis]. This begs the question against SA (which requires no such connection between antecedent and consequent), but in any case let us suppose he is right. It is not my position to defend SA in light of intuitively false counterpossibles like (M). There is still a problem with the reasons he cites for thinking (M) false. First, Goodman cannot rely on ordinary language considerations for ruling out the truth of (M) since they would also rule out of the truth of (N) Were Sylvan’s box empty and non-empty, swans would be red, which he thinks is true because all counterlogicals are true (on which more shortly). Second, nor can he rely on lack of connection between antecedent and consequent as grounding the falsity of (M), for there is equally no connection between antecedent and consequent in (N) which, again, he thinks is true. Why does he think (N) is true in the first place? If he thought it false for the same reasons as (M), there would be no problem (save spelling out the notion of connectedness) in rejecting the truth of (M) because of the lack of connection between antecedent and consequent.

He gives two reasons for thinking (N) true. The first is what he calls the “shrug defense”, originally endorsed by Lewis regarding the vacuous truth of counterpossibles: “[c]onfronted by an antecedent that is not really an entertainable supposition, one may react by saying, with a shrug: If that were so, then anything you like would be true!” [Lew73, p. 24] What is important, however, is that Lewis invokes the shrug defense for the general claim that every counterpossible is true. Goodman, on the other hand, invokes the same defense for a version of that claim restricted to counterlogicals. There are two problems with this. First, reasons given for claims of an unrestricted sort very rarely apply to restricted versions of those claims, e.g. think of Lewis’ reasons for endorsing unrestricted mereological summation—that there is no principled way of restricting it! Second, the restricted version

\textsuperscript{3}All references to Stalnaker are to [Sta68].

\textsuperscript{4}Stalnaker’s semantics includes exactly one impossible world which Goodman calls the ‘Absurd World’, the world at which every sentence is true. A counterpossible is true just in case its consequent holds at the Absurd World. As every sentence holds there, every counterpossible is vacuously true. An important point, then, is that the inclusion of impossible worlds into the semantics does not by itself provide non-vacuous truth conditions for counterpossibles.
of the claim is even less plausible than the unrestricted claim, since no principled reason for the restriction is forthcoming. (Recall Goodman’s reason for rejecting the unrestricted claim—viz., that sentences like (M) are false because of lack of connection between antecedent and consequent. Well the same is true of (N), so that can’t be a reason for the restriction.) Thus the shrug defense used to endorse a less plausible claim will be even less compelling.  

The second reason Goodman gives for thinking (N) true is that he takes the new problem of counterpossibles “to be a sound argument against the best attempts to account for the truth-values of counterlogicals” [Goo04, p. 52]. But he takes the soundness of the argument to rest on the claim that all counterlogicals are true! The circularity is vicious. However, whether all counterlogicals are true is beside the point and the new problem doesn’t rely on it. What is essential to the new problem are the following two assumptions: (i) that certain counterpossibles are intuitively false (so as to motivate the inclusion of non-trivial impossible worlds into the semantics), and (ii) that certain counterlogicals are intuitively true (so as to show these semantics thus extended do not provide the right truth conditions). But the counterlogicals needed to be true are not those like (N). Indeed, the counterlogicals that Goodman considers are the following:

(CP1) If snow were white and non-white, then snow would be white, and
(CP2) If snow were white and non-white, then snow would be non-white.

Notice that each passes the “connection test”—their antecedents and consequents are connected—and intuitively each is true. That’s enough to endorse a disparity in truth values between certain counterlogicals and certain counterpossibles for getting the new problem off the ground. I suspect the reason Goodman endorsed the claim that every counterlogical is true is that this general claim guarantees the truth of both (CP1) and (CP2). But there is no requirement that the intuitive truth of these statements need follow from some more general claim about the truth of all counterlogicals. In the next section I present the new problem for counterpossibles and argue that it goes wrong.

---

5 There is further dialectical tension here. Those who deny the vacuous truth of (M) will not find Lewis’ shrug defense convincing—rather they will find it question begging. For if (M) is intuitively false we cannot simply shrug our shoulders and say, “well I guess it’s true after all, since that is what Lewis’ semantics predicts”. As Goodman is one of those who grants the falsity of (M), he should find the shrug defense unconvincing. Instead he endorses it for an arguably less plausible claim than the one for which Lewis endorsed it.

6 Notice that the falsity of (M) follows from the general principle that any counterfactual which fails the connection test is false. We have seen, however, that this principle is incompatible with the truth of (N). Similarly, one might seek to explain in general terms why both (CP1) and (CP2) are true, and Goodman does, but as made clear below, that
3 The new problem of counterpossibles

So let us suppose the truth of (CP1) and (CP2) and the falsity of (M). One way of avoiding the vacuous truth of sentences like (M) is to extend SA to the Extended Standard Account (ESA):

**Extended Standard Account:** A counterfactual, i.e., a sentence R of the form ‘If A were true, then B would be true’, is true [at] a world β iff in the closest possible world or impossible world α to β [where A is true], B is also true; otherwise R is false in β.7

If the closest impossible world makes ‘Saul Kripke is an earthworm’ true and ‘The square root of 9 is 12’ false, as the proponent of ESA presumes, then (M) is false according to ESA.8 So we have an account that gets the intuitive truth conditions of (M) right.

Both SA and ESA make a Stalnakerian uniqueness assumption that for any counterfactual and world w, there is a closest, and hence unique, antecedent-world w’ to w, assuming the antecedent is possible. The analog for Lewis is the so-called ‘limit assumption’ together with the further assumption that all worlds in the inner-most antecedent-permitting sphere agree on the consequent.9 Since Lewis rejects both of these assumptions, his versions of SA and ESA that drop the analog of uniqueness, which Goodman calls ‘Lewis’s Version of SA’ (‘LVSA’) and ‘Lewis’s version of ESA’ (‘LVESA’) respectively, are as follows:

**LVSA:** A counterfactual, i.e., a sentence R of the form ‘If A were true, then B would be true’, is true at a world β iff either A is impossible, or there is some possible world α where A is true and B is true that is closer to β than any possible world π where A is true and B is false; otherwise R is false in β.

**LVESA:** A counterfactual, i.e., a sentence R of the form ‘If A were true, then B would be true’, is true at a world β iff there is some possible explanation fails. I suggest another explanation in §6 in accord with Goodman’s holding all counterlogicals to be true, but there are others that would deem (CP1) and (CP2) true while still requiring conditionals to pass the connection test (see fn. 14).

7In [Goo04, p. 44], ‘where A is true’ is accidentally left out. I have inserted ‘at’ where ‘in’ occurs since I presume, again, this is a typographical error.

8The details of the account of worlds, possible or not, is unimportant here. All we require of the possible ones is that they be L-maximally classically consistent for the intended language L (say, idealized English) and that the impossible ones be L-maximal. A world w is L-maximal just in case for every sentence A ∈ L, either A is true at w or its negation ¬A is (or both are). A world w is classically consistent just in case for no sentence A ∈ L are both A and ¬A true at w.

9The limit assumption is that any system of spheres be well-ordered by inclusion. This assumption alone doesn’t give us conditional excluded middle on Lewis’ semantics as uniqueness does on Stalnaker’s. Hence the further assumption.
or impossible world $\alpha$ where $A$ is true and $B$ is true that is closer to $\beta$ than any world $\pi$ where $A$ is true and $B$ is false; otherwise $R$ is false in $\beta$.10

The new problem of counterpossibles is that, when we extend SA to ESA and LVSA to LVESA in providing non-vacuous truth conditions for counterpossibles like (M), we wind up incorrectly making at least one of (CP1) and (CP2) false. For, the argument goes, according to ESA the closest impossible world makes either ‘Snow is white’ true and ‘Snow is non-white’ false, or vice versa. If the former, (CP2) is false and if the latter, (CP1) is false. On the other hand, according to LVESA there is no reason to think any ‘Snow is white and non-white’-world where ‘Snow is white’ is true but ‘Snow is non-white’ is not is closer than any ‘Snow is white and non-white’-world where ‘Snow is non-white’ is true while ‘Snow is white’ is not. Hence LVESA deems both (CP1) and (CP2) false.

4 Nearness and ESA

Let ‘Snow is white and non-white’ be denoted by $\lambda$, ‘Snow is white’ by $S$ and ‘Snow is non-white’ by $\neg S$. Then since worlds are well-ordered according to ESA (relative to actuality and a given proposition, in this case $\lambda$), there is a closest $\lambda$-world $w$ which must be one of the following three worlds:

1. $T_1$, at which $S$ but not $\neg S$ is true;
2. $T_2$, at which $\neg S$ but not $S$ is true;
3. $H$, at which both $S$ and $\neg S$ are true.

In presenting the new problem above, the third option, where $w$ is $H$, was ignored. Goodman considers it toward the end as a possible rejoinder to the new problem but dismisses it. Indeed he must dismiss it, since if $w$ is $H$, then both (CP1) and (CP2) are true and ESA gives the correct verdict. But on what grounds can we dismiss that $w$ is $H$? After all, $H$ is the only world of the three that shares with our world the logical law of Simplification, i.e. from $A \land B$ infer both $A$ and $B$, and if match of logical laws bears significantly on nearness like match of natural laws does for Lewis [Lew79, p. 472], then that $w$ is $H$ is a live option.

However, Goodman argues that $T_1$ is closer to the actual world than either $T_2$ or $H$. He explains:

\[\text{Goodman explicitly gives distinct Stalnakerian formulations of SA and ESA, which he labels ‘Stalnaker’s Version of SA’ (‘SVSA’) and ‘Extension of SVSA’ (‘SVESA’), but they’re equivalent to SA and ESA so I will use just ‘SA’ and ‘ESA’ to refer to these accounts.}\]
not only is it true here in our world that \([S]\), but it is false in our world that \([\neg S]\). It is not false in the [world where the third option is satisfied]...that world, unlike actuality, judges that proposition to be true. The extended theorist ought to value this match of particular fact with our world...and hold that these considerations together outweigh the logical similarity of simplification. [Goo04, p. 60]

I agree that match of particular fact will have to outweigh match of logical laws if one wishes to avoid the vacuous truth of all counterlogicals. For if match of logical laws trumps match of particular fact, the Absurd World will always be the closest counterlogical world to actuality since it is the only counterlogical world where all the laws of (classical) logic hold.\(^{11}\) Nonetheless, there is a problem with the quoted passage concerning what is meant by ‘false’. Typically the falsity of \(A\) is defined to be the truth of \(\neg A\), and as such the falsity of \(\neg S\) amounts to nothing more than the truth of \(S\). But then \(\neg S\) is false at both H (and T2), contra Goodman. It just so happens that \(\neg S\) is also true at H! And that’s the whole point of impossible worlds: they are ones at which sentences may be both true and false. After all, \(\neg\) is supposed to be linked to falsity in precisely the following way: \(A\) is false iff \(\neg A\) is true and \(A\) is true iff \(\neg A\) is false, regardless of whether falsity is defined at all. So phrasing the difference between actuality and H concerning \(\neg S\) cannot be done in terms of falsity, for both of them deem \(\neg S\) false.

However the difference could be rephrased in terms of untruth rather than falsity: T1 is closer than H because, like the actual world, \(\neg S\) is untrue there. Now we have a real difference between T1 and H. But it is not clear that match between actuality and T1 concerning the untruth of \(\neg S\) outweighs match of Simplification (with regard to \(\lambda\)) between actuality and H. If we already have a match regarding the truth and falsity of \(S\) and \(\neg S\) between actuality and H, just how much nearer to actuality does match of untruth concerning \(\neg S\) get us over match of Simplification? I would think none for two reasons. First, *in actuality there is no difference between untruth and falsity* (since a proposition is actually untrue iff it’s actually false), so from the point of view of actuality, both H and T1 match with respect to the assignment of truth values to \(S\) and \(\neg S\). Maybe the fact that untruth and falsity are actually coextensive has little to no bearing on whether H is nearer than T1, for we might think the assessment of nearness is not something that is world-relative. Even so, there is a much more significant reason for thinking that match of untruth with regard to \(\neg S\) has no bearing on nearness in the case at hand and that, in fact, we should expect mismatch.

\(^{11}\)With that said, I think Goodman should favor match of logical laws over match of particular fact for the reasons discussed in §6.
When we evaluate a counterfactual with an (actually) false antecedent, we consider only worlds where the antecedent is true. It is crucial, then, that the counterfactual worlds under consideration in the evaluation of a counterfactual conditional (with false antecedent) diverge from actuality as regards the truth value of the antecedent. Now consider the following highly plausible principle:

- In evaluating a counterfactual of the form ‘If it were the case that \(A \land B\), it would be the case that \(C\)” where \(A\) is true and \(B\) is false, we consider only those counterfactual worlds where both \(A\) and \(B\) are true.

Why not require just \(B\) be true and permit \(A\)’s being false? Because worlds at which \(A\) is false and \(B\) is true deviate more from actuality, and hence are further away from actuality, than those at which \(A\) is true and \(B\) is false. For example, in evaluating ‘If it were the case that Caesar crossed the Rubicon and the polar ice caps melted, Northern Canada would be flooded’, we look at only the worlds where Caesar crossed the Rubicon and the polar ice caps melted because they are closer to actuality than those at which Caesar didn’t cross the Rubicon and the polar ice caps melted. Now consider (CP1) and (CP2) again. Since the antecedent \(\lambda\) is a conjunction one of whose conjuncts \(S\) is true and the other \(\neg S\) false, in evaluating either of (CP1) or (CP2) we consider—in accordance with the above principle—only those worlds at which both \(S\) and \(\neg S\) are true, i.e. we consider only \(H\). Whence both (CP1) and (CP2) are true according to ESA after all.

5 Nearness and LVESA

Goodman’s argument above that T1 is closest to actuality is conditional on Stalnaker’s uniqueness assumption (that there is always a closest antecedent-world) which is required by ESA. However, he accepts Lewis’ arguments against uniqueness, and hence favors a Lewisian style account for counterfactuals, and maintains that all worlds where explicit contradictions are true (“counterlogical worlds”) are equidistant from actuality (see [Goo04, p. 58, fn. 46]). The problem is that, if all counterlogical worlds are equidistant from actuality then (CP1) and (CP2) are—contra Goodman—both false according to LVESA. Why then does Goodman at the same time contend they are both true? It appears there is a conflict between the truth values Goodman intuitively assigns to (CP1) and (CP2) (i.e. truth) and the truth values that would be assigned to them on his preferred semantics for counterfactuals which he apparently does not realize is falsity.

So Goodman must reject the equidistance assumption if he wishes, as he indeed does, to hang onto a semantics like LVESA.\(^{12}\) But as soon as

\(^{12}\)See [Goo04, pp. 63–4] for an endorsement of a Lewisian/Stalnaker-type account of
equidistance is given up, there’s no longer any reason to think that the inner-most sphere at which $\lambda$ is entertainable is not one in which the material conditionals $\lambda \rightarrow S$ and $\lambda \rightarrow \neg S$ hold throughout, e.g. because it contains only worlds like H. But that is just to say that both (CP1) and (CP2) may be true according to LVESA after all.

6 Conclusion

Holding on to ESA or LVESA while simultaneously granting the truth of both (CP1) and (CP2) requires one to make crucial choices regarding the nearness of impossible worlds. In particular, it plausibly requires that match of logical laws like Simplification trumps match of particular facts like a sentence’s being untrue (rather than false). But this looks like a choice Goodman could be happy with, for if match of logical laws trumps match of particular facts, then there is good reason to think that the nearest counterlogical world is always the Absurd World, and hence every counterlogical, including (CP1) and (CP2), is true just as Goodman would have it. Of course we would then have to qualify the explanation for (M)’s being false: connection between antecedent and consequent matters only when the antecedent is logically possible. More would need to be said in defense of this qualification, however, though I suspect Goodman would just give a shrug.

The trumping power of match of logical laws implies nothing about the truth of counterpossibles in general, so one is free to hold the vacuous truth of counterlogicals while simultaneously holding that some counterpossibles like (M) are false. Thus an LVESA or ESA account together with a nearness principle stating that match of logical laws trumps match of particular facts (and match of metaphysical laws), gives an account of counterpossibles according to which all counterlogicals are true while some counterpossibles may be false, and this is just the kind of account Goodman is seeking.

Of course we would then have to qualify the explanation for (M)’s being false: connection between antecedent and consequent matters only when the antecedent is logically possible. More would need to be said in defense of this qualification, however, though I suspect Goodman would just give a shrug.

To my mind, the very same considerations for thinking (M) false, viz. lack of connection between antecedent and consequent, apply equally to counterlogicals like (N). Hence my preferred account of counterfactuals is one according to which ‘If it were the case that $A$, it would be the case that $B$’ is true just in case in the inner-most sphere at which $A$ is entertainable, the $R$-relevant (rather than material) conditional $A \rightarrow B$ holds throughout. Cf. [MF95].
References


