On Counterpossibles*

Jens Christian Bjerring

Forthcoming in *Philosophical Studies*

Abstract

The traditional Lewis-Stalnaker semantics treats all counterfactuals with an impossible antecedent as trivially or vacuously true. Many have regarded this as a serious defect of the semantics. For intuitively, it seems, counterfactuals with impossible antecedents—counterpossibles—can be non-trivially true and non-trivially false. Whereas the counterpossible "If Hobbes had squared the circle, then the mathematical community at the time would have been surprised" seems true, "If Hobbes had squared the circle, then sick children in the mountains of Afghanistan at the time would have been thrilled" seems false.

Many have proposed to extend the Lewis-Stalnaker semantics with *impossible worlds* to make room for a non-trivial or non-vacuous treatment of counterpossibles. Roughly, on the extended Lewis-Stalnaker semantics, we evaluate a counterfactual of the form "If A had been true, then C would have been true" by going to closest world—whether possible or impossible—in which A is true and check whether C is also true in that world. If the answer is "yes", the counterfactual is true; otherwise it is false. Since there are impossible worlds in which the mathematically impossible happens, there are impossible worlds in which Hobbes manages to square the circle. And intuitively, in the closest such impossible worlds, sick children in the mountains of Afghanistan are not thrilled—they remain sick and unmoved by the mathematical developments in Europe. If so, the counterpossible "If Hobbes had squared the circle, then sick children in the mountains of Afghanistan at the time would have been thrilled" comes out false, as desired.

In this paper, I will critically investigate the extended Lewis-Stalnaker semantics for counterpossibles. I will argue that the standard version of the extended semantics, in which impossible worlds correspond to maximal, logically inconsistent entities, fails to give the correct semantic verdicts for many

*I am grateful to Daniel Nolan, Weng Hong Tang, and an anonymous referee for *Philosophical Studies* for extremely helpful comments on an earlier draft of this paper. Thanks also to Jørgen Berthelsen, Eline Busck Gundersen, Asbjørn Steglich-Petersen, and audience in the Research Unit for Epistemology, Metaphysics, and Philosophy of Cognition at Aarhus University.

counterpossibles. In light of the negative arguments, I will then outline a new version of the extended Lewis-Stalnaker semantics that can avoid these problems.

1 Introduction

If Timonthy Williamson is right, philosophical reasoning is at its core *counterfactual reasoning* (Williamson 2007). But much useful philosophical reasoning, it seems, also involves *counterpossibles*: counterfactuals with impossible antecedents. We use counterpossibles to convey philosophical information. Logic teachers, for instance, inform students about various non-classical logics by saying things like "If intuitionistic logic were correct, then the law of excluded middle would fail" and "If paraconsistent logic had been true, then logical consequence would not have been explosive"—*psst*, neither intuitionistic nor paraconsistent logic can possibly be true. We also argue philosophically by employing counterpossibles. Philosophy teachers, for instance, may try and convince their students about the untenability of certain ideas by saying things like "Look, if there were a recursive computer that could prove any mathematical sentence that is true, then Gödel's incompleteness theorem would be false. Hence there cannot be such a computer"—*psst*, Gödel's incompleteness theorem cannot possibly be false.

We also seem to have solid semantic intuitions about counterpossibles—assuming, as most others working on counterpossibles, that classical logic is indeed the *one true logic*. The counterpossible "If classical logic had been false, then disjunctive syllogism would no longer hold" seems non-trivially *false*. Perhaps intuitionistic logic had been correct, had classical logic not, and in intuitionistic logic, disjunctive syllogism remains valid. In contrast, the counterpossible "If intuitionistic logic had been correct, then the law of excluded middle would not be unrestrictedly valid" seems non-trivially *true*. For given sufficient knowledge about the formal properties of intuitionistic logic, we know that the law of excluded middle fails in that logic.

Counterpossibles are a species of counterfactuals. Roughly, according to the standard Lewis-Stalnaker semantics, a counterfactual $A \mathrel{\Box\!\!\rightarrow} C$ is true just in case C is true in the closest or most similar possible worlds to the actual world in which A is true.[1] Notoriously, however, this semantics deems all counterpossibles trivially or vacuously *true*. If A is impossible, there are no possible worlds in which A is true, and by convention, $A \mathrel{\Box\!\!\rightarrow} C$ comes out trivially true. But if at least some counterpossibles can be non-trivially false, and some non-trivially true, as the considerations above

[1]For canonical statements of the semantics for counterfactuals, see Lewis 1973 and Stalnaker 1968.

suggest, then the Lewis-Stalnaker semantics seems inadequate.

Many have proposed to *extend* the Lewis-Stalnaker semantics with *impossible worlds* to make room for a non-trivial or non-vacuous treatment of counterpossibles. Roughly, according to the extended Lewis-Stalnaker semantics, a counterfactual $A \mathrel{\Box\!\!\rightarrow} C$ is true just in case C is true in the closest or most similar worlds—whether possible or impossible—to the actual world in which A is true. Although I will be more precise in section 2, we can generally think of impossible worlds as worlds in which the *a priori impossible* happens: worlds in which conceptual, mathematical, and logical truths are false, and in which conceptual, mathematical, and logical falsehoods are true

Consider "If there were a recursive computer that could prove any mathematical sentence that is true, then Gödel's incompleteness theorem would be true." To evaluate this counterpossible, according to the extended Lewis-Stalnaker semantics, we go to the closest or most similar impossible world w in which there is a recursive computer that can prove any true mathematical sentence and ask: is Gödel's theorem true in w? Presupposing an intuitive grasp of relative closeness or similarity between worlds, we should answer "no" to this question: if there were a recursive computer that could prove any true mathematical sentence, Gödel's theorem would be false precisely because it rules out the possibility of such a computer. If so, the extended Lewis-Stalnaker semantics tells us that the counterpossible is false. And intuitively, it *is* false, albeit only non-trivially so. Consider also "If intuitionistic logic were correct, then the law of excluded middle would fail." To evaluate this counterpossible, we go to the closest or most similar impossible world w in which intuitionistic logic is correct and ask: does the law of excluded middle fail in w? Presupposing an intuitive grasp of relative closeness or similarity between worlds, we should answer "yes" to this question: if intuitionistic logic were indeed true, there would be counterexamples to the law of excluded middle—at least in infinite domains. If so, the extended Lewis-Stalnaker semantics tells us that the counterpossible is true. And intuitively, it *is* true, albeit only non-trivially so.

The extended Lewis-Stalnaker semantics is arguably also the standard account of counterpossibles in the literature, and it has been endorsed by Berit Brogaard & Joseph Salerno, Daniel Nolan, and David Vander Laan amongst others. In this paper, however, I will argue that the extended Lewis-Stalnaker semantics fails when impossible worlds correspond to *maximal*, logically inconsistent entities for which it holds that either A or $\neg A$ is true, for every sentence A. This negative result will be particularly troublesome for Brogaard & Salerno 2013 and Vander Laan 2004 who explicitly conceive of impossible worlds as maximal, inconsistent sets of sentences or propositions—see also Jeffrey Goodman 2004 and Edwin Mares 1997.

As far as I can tell, there are two ways in which a proponent of the extended Lewis-Stalnaker semantics may try to avoid my negative result. On the one hand, she may follow Nolan 1997 and appeal to *non-maximal* or *partial* impossible worlds in which sentences may fail to receive a truth-value. Curiously, Vander Laan notices that his account of counterpossibles—in which all impossible worlds are maximal—"differs *slightly* from Nolan [1997], in which worlds may have truth value gaps." (Vander Laan 2004, p. 259; my italics.) If I am right, however, this difference is not slight. Rather, it could be mean the difference between failure and potential success for the extended Lewis-Stalnaker semantics. But if so, as I argue in section 4.1, there should be a greater focus on how exactly to understand non-maximal worlds and their role in a semantics for counterpossibles. On the other hand, as I discuss in section 4.2, a proponent of the extended Lewis-Stalnaker semantics may try to avoid the negative result by identifying impossible worlds with linguistic entities that are closed under logical consequence in some non-classical logic. I will devote most of section 4 to this second option. Not only because I find this version of the extended Lewis-Stalnaker semantics the most promising, but also because it constitutes the outlines of a novel semantics for counterpossibles.

I proceed as follows. In section 2, I introduce the extended Lewis-Stalnaker semantics in more details. In particular, I spell out what I—alongside many others—take logically possible and logically impossible worlds to be. In section 3, I show that the extended Lewis-Stalnaker semantics fails when impossible worlds correspond to maximal, logically inconsistent entities. In section 4, I discuss how a proponent of the extended semantics may attempt to avoid the negative result. In section 5, I conclude.

2 Counterpossibles, semantics, and world-ontology

Let us first state more precisely the standard Lewis-Stalnaker semantics for counterfactuals:

> (**LS-ConFac**) $A \mathrel{\Box\!\!\rightarrow} C$ is true in the actual world w_α iff some possible world in which A and C are true is closer to w_α than any possible world in which A is

true and C is false.

While (LS-ConFac) captures the general idea behind the standard Lewis-Stalnaker semantics, it is specific in the following two respects. First, counterfactuals are only evaluated for truth and falsity in the actual world. This restriction will simplify the presentation of the *extended* Lewis-Stalnaker semantics, which involves reference to impossible worlds. For while it is quite clear how to evaluate counterfactuals in arbitrary possible worlds, it is not so clear what it means to evaluate counterfactuals in arbitrary impossible worlds. Second, (LS-ConFac) favors the Lewisian semantics for counterfactuals over the Stalnakerian one in the sense that it validates neither the *uniqueness* nor the *limit* assumption. Against the uniqueness assumption, that is, worlds can tie for closeness to w_α, and against the limit assumption, worlds can be closer and closer to w_α without end.[2]

When A is impossible, (LS-ConFac) always delivers the value "true" for $A \Box\!\!\rightarrow C$. I will assume that counterpossibles trade non-trivially in the impossible, and that (LS-ConFac) is inadequate because it vacuously deems all counterpossibles true. Both these assumptions can of course be disputed, but in this paper I shall focus solely on the extended Lewis-Stalnaker semantics for counterfactuals:

> (**ConFac**) $A \Box\!\!\rightarrow C$ is true in the actual world w_α iff some world in which A and C are true is closer to w_α than any world in which A is true and C is false.

In contrast to the world-ontology underlying (LS-ConFac), the one underlying (ConFac) contains both possible and impossible worlds. While we can assume that (LS-ConFac) and (ConFac) agree on the evaluation of counterfactuals, they need not agree about the evaluation of some counterpossibles. For since "worlds" in (ConFac) may refer both to possible and impossible worlds, we can use the semantics, as illustrated above, to deem certain counterpossibles non-trivially false and others non-trivially true. In (ConFac), relative closeness may obtain not only between possible worlds but also between the actual world and impossible worlds. A lot can be said about relative closeness between possible and impossible worlds, but luckily, my central arguments remain largely unaffected by the finer details of how to understand closeness.

We can now start to look more carefully at the world-ontology underlying (ConFac). Throughout the paper, I will identify both possible and impossible worlds with

[2] Also, I have for simplicity omitted an accessibility relation between worlds in (LS-ConFac).

sets of sentences in some sufficiently strong world-making language. We already have a good understanding of what it means to give an ersatz or a linguistic construction of possible worlds as sets of sentences in some world-making language, and we will be able to use the same tools to construct impossible worlds.[3] Obviously, there are tricky questions about what a sufficiently powerful world-making language should look like, but for current purposes, I will simply assume that we have an account of such a language.[4] The language, however, is stipulated to have the following two features. First, it contains all (possible) sentence types of English—or sufficiently regimented such sentence types. This will facilitate the accounts below of what it means for a sentence to be true or false in a world. Second, the language contains symbols $\neg$ and $\wedge$ that play the same inferential roles as classical negation and conjunction. The other standard connectives of propositional logic can then be treated as shorthand for their definitions in terms of $\neg$ and $\wedge$.[5]

My discussion will center around the class of counterpossibles whose antecedents presuppose that classical (propositional) logic is false, or that some non-classical logic is true. I will also call such counterpossibles *counterlogicals*. Given that worlds correspond to sets of sentences (in the assumed world-making language), we can then begin to lay down principles that distinguish *logically possible* from *logically impossible* worlds. Following Francesco Berto 2009, there are generally two ways in which we can think of (logically) impossible worlds.

We can think of impossible worlds as "American type" impossible worlds. These correspond to sets of sentences that are (classically) logically inconsistent. Arguably, the standard extended Lewis-Stalnaker semantics presupposes that impossible worlds are American type impossible worlds, and, in particular—modeled on the familiar conception of logically possible worlds as *maximal, consistent sets of sentences*—that impossible worlds correspond to *maximal*, inconsistent sets of sentences. While Brogaard & Salerno 2013, Goodman 2004, and Vander Laan 2004 all share this conception

[3] For various linguistic constructions of possible worlds, see for instance Robert Adams 1974, Rudolf Carnap 1947, David J. Chalmers 2011, Jaakko Hintikka 1969, and Richard Jeffrey 1983.

[4] For further discussion of what a general purpose world-making language may look like, refer to Chalmers 2011 and Lewis 1986, chapter 3. Insofar as one has a better grasp of a class of primitive, arbitrarily fine-grained propositions that can stand in logical relations to each other, it is also worth pointing out that one might replace all talk about sentences with talk of such propositions without affecting the main ideas and arguments in the paper.

[5] Although I will work with a simple logical language that contains only $\neg$ and $\wedge$ as logical symbols, it will be easy to see that the results in section 3 will continue to hold if we enrich the world-making language to include symbols for the other propositional connectives as well.

of impossible worlds, Nolan 1997 allows that some American type impossible worlds are also non-maximal. Alternatively, we can think of impossible worlds as "Australian type" impossible worlds. These correspond to sets of sentences that are closed under logical consequence in some non-classical logic. Edwin Mares 2004 appeals to Australian types impossible worlds in his semantics for counterpossibles.

In what follows, I will first assume that all impossible worlds in the ontology underlying (ConFac) are American type impossible worlds. In particular, I will assume that while logically possible worlds correspond to maximal, consistent sets of sentences, logically impossible worlds correspond to maximal, inconsistent sets of sentences. In section 4, I will then look at what happens if we give up the requirement that all American type impossible worlds are maximal, and, in more detail, at what happens if we think of impossible worlds as Australian type impossible worlds.

2.1 Worlds: possible and impossible

We want to develop a world-ontology in which possible worlds correspond to maximal, consistent sets of sentences, and in which impossible worlds correspond to maximal, inconsistent such sets. Let us first give a simple account of what it means for a sentence to be true in a world:

(**Truth**) A sentence A is *true* in a world w iff $A \in w$.
(**Falsity**) A sentence A is *false* in a world w iff $A \notin w$.

When A is true in w, I will also say that *w verifies A*. While Brogaard & Salerno 2013 gives a similar account of truth and falsity in a world, notice that my central arguments in section 3 do not hang on this simple account. Instead we may spell out falsity of A in w in terms of truth of $\neg A$ in w, or we may simply take truth-in-a-world as a primitive notion.

We can now define what it means for a set of sentences to be logically consistent. Since I can make my main arguments by focusing on propositional logic, I will focus on logical consistency in propositional logic:

> (**Consistency**) A set Γ of sentences is *consistent* iff Γ is satisfiable, where Γ is *satisfiable* iff there is a propositional evaluation that makes all sentences in Γ true.

Let $\mathcal{I}$ be an interpretation function that assigns either true or false, but not both, to each atomic sentence A and define ν to be the following classical evaluation function:

(**Evaluation**)

($\nu\mathcal{I}$) If A is an atomic sentence, then $\nu(A) = \mathcal{I}(A)$.

($\nu\neg$) $\nu(\neg A) = T$ iff $\nu(A) = F$.
($\nu\wedge$) $\nu(A \wedge B) = T$ iff $\nu(A) = T$ and $\nu(B) = T$.

The semantic clauses for $\neg$ and $\wedge$ reflect the idea that the world-making language has symbols that play the same inferential roles as classical negation and conjunction. Since we can express sentences involving the other standard propositional connectives using just $\neg$ and $\wedge$, ν governs indirectly all these sentences as well.

Given this, (Satisfiability) can be proved:[6]

> (**Satisfiability**) Any set Γ of sentences that satisfies the following two conditions is satisfiable:
> (i) $A \in \Gamma$ iff $\neg A \notin \Gamma$.
> (ii) $(A \wedge B) \in \Gamma$ iff $A \in \Gamma$ and $B \in \Gamma$.

We can then say that every set of sentences—and hence every world—that satisfies (i) and (ii) is consistent, and that every inconsistent set of sentences fails to satisfy either (i) or (ii).

Finally, we can define what it means for a set of sentences to be maximal:

> (**Maximality**) A set Γ of sentences is *maximal* iff for all sentences A, either $A \in \Gamma$ or $\neg A \in \Gamma$.

If a set of sentences—and hence a world—does not satisfy (Maximality), I will also say that the set is *non-maximal* or *partial*.

The following two principles now hold:

> (**Ent**$_1$) For all sentences A and B such that A logically entails B, if A is true in w, then B is true in w.
>
> (**Ent**$_2$) For any tautology T, T is true in w.

(Ent$_1$) follows from the fact that for any maximal, consistent set Γ of sentences, and for any sentence B such that Γ logically entails B, $B \in \Gamma$. And (Ent$_2$) follows from the fact that any set of sentences logically entails any tautology T. Given (Ent$_1$)

[6]See the appendix for the proof of (Satisfiability).

and (Ent_2), maximal consistent sets of sentences can hence play the role of logically possible worlds. For logically possible worlds are the kinds of entities that verify all logical consequences of sentences that they already verify—including all logical truths.[7]

Given that logical possible worlds correspond to maximal, consistent sets of sentences, we can then identify logically impossible worlds with maximal sets of sentences that are inconsistent. In turn, we can identify the world-ontology underlying (ConFac) with the class of logically possible and impossible worlds so construed. Call this version of (ConFac) in which all worlds are maximal "(M-ConFac)".

In arguing against (M-ConFac) below, I will assume—as most others endorsing the extended Lewis-Stalnaker semantics—that the closest worlds relevant for evaluating counterlogicals are always also logically impossible. In determining what would have followed, for instance, had some non-classical logic such as intuitionistic logic been correct, we consider worlds whose truths are *not* codified by classical logic but rather by the non-classical logic in question. I stress this point because it might be argued that we can handle certain counterlogicals such as "If intuitionistic logic had been correct, then the law of excluded middle would not be unrestrictedly valid" without appealing to *logically* impossible worlds in the sense above.[8] Consider a world w pretty much like ours but in which the sentences "Intuitionistic logic is the correct logic" and "The law of excluded middle is not unrestrictedly valid" are true rather than the actually true sentences "Classical logic is the correct logic" and "The law of excluded middle is unrestrictedly valid." Although w is logically consistent, and hence not logically impossible, there is plausibly a sense in which w is impossible nonetheless. If so, we can appeal to w to ensure that the extended Lewis-Stalnaker semantics renders true the counterlogical "If intuitionistic logic had been correct, then the law of excluded middle would not be unrestrictedly valid." So it seems that we need not always appeal to logically impossible worlds to evaluate counterlogicals.

Even if we grant that a world like w is in some sense impossible, it seems nevertheless clear that the *closest* worlds relevant for evaluating counterlogicals are *logically* impossible ones. Consider an analogy. If we are asked to evaluate what would have

[7]Assuming that general-purpose possible worlds are also mathematically and analytically possible, maximal consistent sets of sentences clearly cannot play this general-purpose role. Yet, since my main focus is on counterlogicals, we can set aside issues about non-logical a priori truths such as those resulting from mathematical and analytic reasoning.

[8]Thanks to an anonymous referee for discussion here.

happened, had some actual law of nature $L_{\mathcal{N}}$ been false, we are asked to consider what happens in a physically or nomologically impossible world whose truths and physical features cannot adequately be described by $L_{\mathcal{N}}$. We are not merely asked to consider what happens in a world that verifies the sentence "$L_{\mathcal{N}}$ is false" rather than the sentence "$L_{\mathcal{N}}$ is true", but whose laws of nature are otherwise identical to the ones in our world. Similarly, if we are asked to evaluate what would have happened, had some non-classical logic $\mathcal{L}_i$ been correct, we are asked to consider what happens in a world whose truths and logical features cannot adequately be described or codified by classical logic. We are not merely asked to consider what happens in a world that verifies sentences such as "$\mathcal{L}_i$ is the correct logic" rather than "Classical logic is the correct logic", but whose logical structure otherwise is identical to the one in our world. Rather, we consider what happens in a logically impossible world whose truths are governed by $\mathcal{L}_i$ rather than classical logic. So from now on, I take it, we must appeal to logically impossible worlds to evaluate counterlogicals.[9]

My arguments against (M-ConFac) center around counterlogicals, and, as such, around the use of logically impossible worlds in the extended Lewis-Stalnaker semantics. I will argue that (M-ConFac) fails insofar as it is to play the role of a general semantics that can cover all types of counterpossibles, and notably, counterlogicals. But note that these arguments leave open the possibility that (M-ConFac) can correctly handle certain counterpossibles. On some conceptions of the impossible, worlds in which water is not H_2O are metaphysically impossible, and counterfactuals whose antecedents presuppose that water is not H_2O can in turn be thought of as counterpossibles. Since we need not appeal to logically impossible worlds to account for counterfactuals with metaphysically impossible antecedents, my arguments will not show that (M-ConFac) fails for such counterfactuals. Similarly, insofar as we can handle counterfactuals with analytically and mathematically impossible antecedents without appealing to logically impossible worlds, my arguments will not show that (M-ConFac) fails for such counterfactuals. Rather, my arguments in the next section will show that (M-ConFac) fails as a general semantics for counterpossibles.

[9]For further discussion of logically consistent worlds that represent themselves as being logically non-classical in different ways, see Krakauer 2012.

3 Problems

I shall now argue that (M-ConFac) fails as a general semantics for counterpossibles. Henceforth, I will mean *logically* impossible worlds—in the sense above—whenever I say "impossible worlds" without qualification, and I will mean *classical propositional logic* whenever I use locutions such as "logical entailment", "logical truth", and "logical consistency" without qualification.

My main line of arguing against (M-ConFac) relies on (Inc), which can be proved using (Satisfiability):[10]

> (**Inc**) All maximal, logically inconsistent sets of sentences (in the world-making language) contain an instance of a LNC-, CF-, or NCF-inconsistency:
>
> **LNC-inconsistency** (law of non-contradiction): $\{A, \neg A\}$.
>
> **CF-inconsistency** (conjunction fallacy): $\{\neg A, (A \wedge B)\}$, $\{\neg B, (A \wedge B)\}$, $\{\neg A, \neg B, (A \wedge B)\}$.
>
> **NCF-inconsistency** (negated conjunction fallacy): $\{\neg(A \wedge B), A, B\}$.

Given (Inc), we can divide all *logically* impossible worlds into the following types:

> **Type-1 impossible worlds**: worlds that contain a LNC-inconsistency.
>
> **Type-2 impossible worlds**: worlds that contain a CF-inconsistency.
>
> **Type-3 impossible worlds**: worlds that contain a NCF-inconsistency.

In evaluating a given counterlogical using (M-ConFac), we know then that the relevant antecedent worlds must be either type-1, type-2, or type-3 impossible worlds.

Consider now:

(1) If intuitionistic logic were correct and $(A \wedge B)$ true, then A and B would also be true (for any A and B).

Given that classical logic is in fact the one true logic, intuitionistic logic cannot possibly be true. So (1) is a counterlogical. (1) is also a *true* counterlogical because intuitionistic logic respects the standard logical laws governing conjunctions, and the set $\{(A \wedge B), \neg A, \neg B\}$ is not intuitionistically satisfiable.

[10] See the appendix for the proof of (Inc).

If (M-ConFac) is to be a successful semantics, it must render (1) true. To do so, some logically impossible world w in which intuitionistic logic, $(A \wedge B), A$, and B are all true must be closer to the actual world w_α than any logically impossible worlds in which intuitionistic logic and $(A \wedge B)$ are true, but either A or B false—and hence, by (Maximality), either $\neg A$ or $\neg B$ true. Given (Inc), we know that w must be either a type-1, a type-2, or a type-3 impossible world. I now show that (M-ConFac) must deliver the wrong semantic verdict for some counterpossible if it is to render (1) true. In the first part of the argument, I show that (M-ConFac) fails as an adequate semantics when there is a unique closest impossible world that verifies the antecedent in (1). In the second part, I then show that (M-ConFac) fails when multiple impossible worlds, each verifying the antecedent in (1), may count as equally close to actuality.

First part

Suppose first that w is a type-1 impossible world in which intuitionistic logic, $(A \wedge B), A$, and B are all true. Suppose also that w is closer to w_α than any impossible world in which intuitionistic logic and $(A \wedge B)$ are true, but either A or B false. In this case, (M-ConFac) will deem (1) true. Since w is a type-1 world, however, we also know that C and $\neg C$ are true in w, for some C. So (M-ConFac) also wrongly deems the following counterpossible true:

(2) If intuitionistic logic were correct and $(A \wedge B)$ true, then C and $\neg C$ would both be true (for some C).

If the closest w is a type-1 world in which C and $\neg C$ are true in addition to intuitionistic logic, $(A \wedge B)$, A, and B, then (2) is true according to (M-ConFac). But (2) should be *false*. For intuitionistic does not tolerate nor license any contradictions, and the set $\{C, \neg C\}$ is not intuitionistically satisfiable for any C. So if (M-ConFac) is to make (1) but not (2) true, w cannot be a type-1 impossible world.

Suppose second that w is a type-2 impossible world in which intuitionistic logic, $(A \wedge B), A$, and B are all true. Suppose also that w is closer to w_α than any impossible world in which intuitionistic logic and $(A \wedge B)$ are true, but either A or B false. In this case, (M-ConFac) will deem (1) true. Since w is a type-2 world, however, we also know that either $(C \wedge D)$ and $\neg C$, $(C \wedge D)$ and $\neg D$, or $(C \wedge D)$, $\neg C$, and $\neg D$ are true in w, for some C and D. So (M-ConFac) also wrongly deems at least one of the following counterpossibles true:

(3) If intuitionistic logic were correct and $(A \wedge B)$ true, then $(C \wedge D)$ and $\neg C$ would all be true (for some C and D).

(4) If intuitionistic logic were correct and $(A \wedge B)$ true, then $(C \wedge D)$ and $\neg D$ would all be true (for some C and D).

(5) If intuitionistic logic were correct and $(A \wedge B)$ true, then $(C \wedge D)$, $\neg C$ and $\neg D$ would all be true (for some C and D).

If the closest w is a type-2 world in which $(C \wedge D)$ and $\neg C$ are true in addition to intuitionistic logic, $(A \wedge B)$, A, and B, then (3) is true according to (M-ConFac). If $\neg D$ is true in w rather than $\neg C$, then (4) is true according to (M-ConFac), whereas if both $\neg C$ and $\neg D$ are true in w, then (5) is true according to (M-ConFac). But (3) to (5) should all be *false*. Intuitionistic logic does not license any violations of the standard laws governing conjunctions, and the sets $\{(C \wedge D), \neg C\}$, $\{(C \wedge D), \neg D\}$, and $\{(C \wedge D), \neg C, \neg D\}$ are not intuitionistically satisfiable. So if (M-ConFac) is to make (1) but not any of (3) to (5) true, w cannot be a type-2 impossible world.

Suppose third that w is a type-3 world in which intuitionistic logic, $(A \wedge B), A$, and B are all true. Suppose also that w is closer to w_α than any impossible world in which intuitionistic logic and $(A \wedge B)$ are true, but either A or B false. In this case, (M-ConFac) will deem (1) true. Since w is a type-3 world, however, we also know that $\neg(C \wedge D)$, C, and D are true in w, for some C and D. So (M-ConFac) also wrongly deems the following counterpossible true:

(6) If intuitionistic logic were correct and $(A \wedge B)$ true, then $\neg(C \wedge D)$, C, and D would all be true (for some C and D).

If the closest w is a type-3 world in which $\neg(C \wedge D)$, C, and D are true in addition to intuitionistic logic, $(A \wedge B)$, A, and B, then (6) is true according to (M-ConFac). But (6) should be *false*. Intuitionistic logic does not license any violations of the standard laws governing conjunctions, and the set $\{\neg(C \wedge D), C, D\}$ is not intuitionistically satisfiable. So if (M-ConFac) is to make (1) but not (6) true, w cannot be a type-3 impossible world.

The discussion above shows that no matter which type of logically impossible world counts as the closest to actuality, (M-ConFac) must deem true some false counterlogicals in order to render (1) true. So whenever there is a unique closest impossible world to w_α, (M-ConFac) cannot be the right semantics for counterpossibles. Put differently: if the semantics satisfies the uniqueness assumption, then (M-ConFac) fails.

To save (M-ConFac) from the problems above, we must ensure that it renders (2) to (6) false in all cases where it renders (1) true. To do so, the uniqueness assumption must go. It is clear that we cannot save (M-ConFac) merely by giving up the limit assumption and allow that some impossible worlds can be closer and closer to actuality without end. Without the limit assumption, (M-ConFac) can render (1) true in the case we never reach a sphere of worlds—as we approach actuality—in which some impossible world verifying intuitionistic logic, $(A \wedge B)$ but not both A and B is closer to actuality than some impossible world verifying intuitionistic logic, $(A \wedge B)$, A, and B. In that case, however, it is easily verified that the problems from above will reoccur for (M-ConFac). For as long as some logically impossible world verifying intuitionistic logic, $(A \wedge B)$, A, and B continues to be closer to actuality than any world verifying intuitionistic logic, $(A \wedge B)$ but not both A and B, we can apply the line of reasoning from above to show that (M-ConFac) must give the wrong semantic verdict for some other counterlogicals. Rather, to ensure that (M-ConFac) renders (2) to (6) false in all cases where it renders (1) true, there must be some logically impossible worlds that always count as equally close to actuality. But even then, as I shall argue now, (M-ConFac) will fail as an adequate semantics.

Second part

To save (M-ConFac) from the problems above, we must assume that there is always at least two relevant type-1, type-2, or type-3 worlds that count as equally close to w_α. Consider first two type-1 worlds w_1 and w_2 in which intuitionistic logic, $(A \wedge B)$, A, and B are true but which differ on which instance of a LNC-inconsistency is true. In particular, while A_1 and $\neg A_1$ are true in w_1, and B_1 and $\neg B_1$ are true in w_2, no other inconsistencies are true in w_1 and w_2 respectively.[11] Suppose that w_1

[11]Strictly, w_1 and w_2 cannot contain just one instance of a LNC-inconsistency, but I simplify the presentation here—as I do below. If need be, we can stipulate that A_1 and B_1 are atomic sentences and that all LNC-inconsistencies in w_1 involve A_1, whereas all LNC-inconsistencies in w_2 involve

and w_2 are closer to w_α than any other impossible world verifying the antecedent in (1), and that w_1 is always as close to w_α as w_2. The counterlogical (1) is now true because the closest impossible worlds to w_α are worlds in which intuitionistic logic, $(A \wedge B)$, A, and B are all true. But at the same time every counterlogical of the following form is *false*:[12]

(7) If intuitionistic logic were correct and $(A \wedge B)$ true, then [instance of LNC-inconsistency] would be true.

(8) If intuitionistic logic were correct and $(A \wedge B)$ true, then [instance of CF-inconsistency] would be true.

(9) If intuitionistic logic were correct and $(A \wedge B)$ true, then [instance of NCF-inconsistency] would be true.

To evaluate (7), we go to the closest impossible worlds in which the antecedent in (7) is true. By assumption, these are w_1 and w_2. Since all LNC-inconsistencies not involving A_1 and B_1 are false in w_1 and w_2, we focus on the pairs $\{A_1, \neg A_1\} \subset w_1$ and $\{B_1, \neg B_1\} \subset w_2$. Since w_1 and w_2 are closest to w_α, worlds verifying intuitionistic logic, A_1, and $\neg A_1$ will *not* be closer to w_α than worlds—namely w_2—verifying intuitionistic logic but not both A_1 and $\neg A_1$. So if we plug the pair $\{A_1, \neg A_1\}$ into the consequent in (7), (7) is false according to (M-ConFac). Similarly, as can be easily verified, (7) is false according to (M-ConFac) if we plug $\{B_1, \neg B_1\}$ into the consequent in (7). Accordingly, (7) is false for any instance of a LNC-inconsistency whenever at least two type-1 worlds count as closest to w_α.

To evaluate (8) and (9), we go to the closest impossible worlds in which the antecedents in (8) and (9) are true. By assumption, these are w_1 and w_2 both of which are only LNC-inconsistent. Hence the consequents in (8) and (9) are false, and in turn (8) and (9) are false for any instance of a CF- and NCF-inconsistency. Since (2) is an instance of (7), (3) to (5) instances of (8), and (6) an instance of (9), (M-ConFac) can render (1) but not (2) to (6) true when two type-1 worlds are closest to w_α.

B_1. Intuitively, on such a picture, if we were to remove all sentences involving A_1 from w_1, and all sentences involving B_1 from w_2, the two worlds would become consistent.

[12] I state (7) to (9) schematically using, for simplicity, expressions from the metalanguage.

So when the two type-1 worlds w_1 and w_2 count as equally close to w_α—nothing hangs on the particular choice of w_1 and w_2—(M-ConFac) deems (1) but not (2) to (6) true. But if it does, it also wrongly deems the following true counterlogicals false:

(10) If intuitionistic logic were correct and $(A \wedge B)$ true, then A_1 and $\neg A_1$ would never both be true.

(11) If intuitionistic logic were correct and $(A \wedge B)$ true, then B_1 and $\neg B_1$ would never both be true.

(10) and (11) are true counterlogicals because intuitionistic logic does not tolerate nor license any contradictions. For (10) to be true, an impossible world verifying intuitionistic logic, $(A \wedge B)$, but not both A_1 and $\neg A_1$ must be closer to w_α than any world verifying intuitionistic logic, $(A \wedge B)$, A_1, and $\neg A_1$. But since w_1 is among the closest worlds to w_α, this condition is not met. So (10) is false. Similarly, for (11) to be true, a world verifying intuitionistic logic, $(A \wedge B)$, but not both B_1 and $\neg B_1$ must be closer to w_α than any world verifying intuitionistic logic, $(A \wedge B)$, B_1, and $\neg B_1$. But since w_2 is among the closest worlds to w_α, this condition is not met. So (11) is false. So if (M-ConFac) is to make (1) but not (2) to (6) true—by means of w_1 and w_2—it fails to make (10) and (11) true. So (M-ConFac) fails when at least two type-1 worlds count as equally close to actuality.

Consider second two type-2 worlds w_3 and w_4 in which intuitionistic logic, $(A \wedge B)$, A, and B are true but which differ on which instance of a CF-inconsistency is true. In particular, while $\neg A_1$, $\neg A_2$, and $(A_1 \wedge A_2)$ are true in w_3, and $\neg B_1$, $\neg B_2$, and $(B_1 \wedge B_2)$ are true in w_4, no other inconsistencies are true in w_3 and w_4 respectively. Suppose that w_3 and w_4 are closer to w_α than any other impossible world verifying the antecedent in (1), and that w_3 is always as close to w_α as w_4. It is then easy to see that (7) and (9)—and hence (2) and (6)—are false while (1) is true. To evaluate (7) and (9), we go to the closest impossible worlds in which the antecedents in (7) and (9) are true. By assumption, these are w_3 and w_4 both of which are only CF-inconsistent. Hence the consequents in (7) and (9) are false. In turn, (7) and (9) are false for any instance of a LNC- and NCF-inconsistency—and along with them (2) and (6)—whenever at least two type-2 worlds are closest to w_α.

To evaluate (8), we go to the closest impossible worlds in which the antecedent in (8) is true. By assumption, these are w_3 and w_4. Since all CF-inconsistencies not involving $\{A_1, A_2\}$ and $\{B_1, B_2\}$ are false in w_3 and w_4, we focus on the sets $\{\neg A_1, \neg A_2, (A_1 \wedge A_2)\} \subset w_3$ and $\{\neg B_1, \neg B_2, (B_1 \wedge B_2)\} \subset w_4$. Since w_3 and w_4 are closest to w_α, worlds verifying intuitionistic logic, $\neg A_1$, $\neg A_2$, and $(A_1 \wedge A_2)$ will *not* be closer to w_α than worlds—namely w_4—verifying intuitionistic logic, A_1, A_2, and $(A_1 \wedge A_2)$. So if we plug the triple $\{\neg A_1, \neg A_2, (A_1 \wedge A_2)\}$ into the consequent in (8), (8) is false according to (M-ConFac). Similarly, as can be easily verified, (8) is false according to (M-ConFac) if we plug $\{\neg B_1, \neg B_2, (B_1 \wedge B_2)\}$ into the consequent in (8). Accordingly, (8) is false for any instance of a CF-inconsistency—and along with it (3) to (5)—whenever at least two type-2 worlds are closest to w_α.

So when the two type-2 worlds w_3 and w_4 count as equally close to w_α—nothing hangs on the particular choice of w_3 and w_4—(M-ConFac) deems (1) but not (2) to (6) true. But if it does, it also wrongly deems the following true counterlogicals false:

(12) If intuitionistic logic were correct and $(A \wedge B)$ true, then $\neg A_1$ and $(A_1 \wedge A_2)$ would never both be true.

(13) If intuitionistic logic were correct and $(A \wedge B)$ true, then $\neg B_1$ and $(B_1 \wedge B_2)$ would never both be true.

(12) and (13) are true counterlogicals because intuitionistic logic does not license any violations of the standard laws governing conjunctions. Since w_3 is among the closest worlds to w_α, however, a world verifying intuitionistic logic but not both $\neg A_1$ and $(A_1 \wedge A_2)$ is *not* closer to w_α than a world—namely w_3—verifying intuitionistic logic, $\neg A_1$, and $(A_1 \wedge A_2)$. So (12) is false. Similarly, since w_4 is among the closest worlds to w_α, a world verifying intuitionistic logic but not both $\neg B_1$ and $(B_1 \wedge B_2)$ is *not* closer to w_α than a world—namely w_4—verifying intuitionistic logic, $\neg B_1$, and $(B_1 \wedge B_2)$. So (13) is false. So if (M-ConFac) is to make (1) but not (2) to (6) true—by means of w_3 and w_4—it fails to make (12) and (13) true. So (M-ConFac) fails when at least two type-2 worlds count as equally close to actuality.

Consider finally two type-3 worlds w_5 and w_6 in which intuitionistic logic, $(A \wedge B)$, A, and B are true but which differ on which instance of a NCF-inconsistency is true. In particular, while A_1, A_2, and $\neg(A_1 \wedge A_2)$ are true in w_5, and B_1, B_2, and $\neg(B_1 \wedge B_2)$ are true in w_6, no other inconsistencies are true in w_5 and w_6 respectively. Suppose that w_5 and w_6 are closer to w_α than any other impossible world verifying the antecedent in (1), and that w_5 is always as close to w_α as w_6. It is then easy to see that (7) and (8)—and hence (1) to (5)—are false while (1) is true. To evaluate (7) and (8), we go to the the closest impossible worlds in which the antecedents in (7) and (8) are true. By assumption, these are w_5 and w_6 both of which are only NCF-inconsistent. Hence the consequents in (7) and (8) are false. In turn, (7) and (8) are false for any instance of a LNC- and CF-inconsistency—and along with them (1) to (5)—whenever at least two type-3 worlds are closest to w_α.

To evaluate (9), we go to the closest impossible worlds in which the antecedent in (9) is true. By assumption, these are w_5 and w_6. Since all NCF-inconsistencies not involving $\{A_1, A_2\}$ and $\{B_1, B_2\}$ are false in w_5 and w_6, we focus on the sets $\{A_1, A_2, \neg(A_1 \wedge A_2)\} \subset w_5$ and $\{B_1, B_2, \neg(B_1 \wedge B_2)\} \subset w_6$. Since w_5 and w_6 are closest to w_α, worlds verifying intuitionistic logic, A_1, A_2, and $\neg(A_1 \wedge A_2)$ will *not* be closer to w_α than worlds—namely w_6—verifying intuitionistic logic, A_1, A_2, and $(A_1 \wedge A_2)$. So if we plug the triple $\{A_1, A_2, \neg(A_1 \wedge A_2)\}$ into the consequent in (9), (9) is false according to (M-ConFac). Similarly, as can be easily verified, (9) is false according to (M-ConFac) if we plug $\{B_1, B_2, \neg(B_1 \wedge B_2)\}$ into the consequent in (9). Thus (9) is false for any instance of a NCF-inconsistency—and along with it (6)—whenever at least two type-3 worlds are closest to w_α.

So when the two type-3 worlds w_5 and w_6 count as equally close to w_α—nothing hangs on the particular choice of w_5 and w_6—(M-ConFac) deems (1) but not (2) to (6) true. But if it does, it also wrongly deems the following true counterlogicals false:

(14) If intuitionistic logic were correct and $(A \wedge B)$ true, then A_1, A_2, and $\neg(A_1 \wedge A_2)$ would never all be true.

(15) If intuitionistic logic were correct and $(A \wedge B)$ true, then B_1, B_2, and $\neg(B_1 \wedge B_2)$ would never all be true.

(14) and (15) are true counterlogicals because intuitionistic logic does not license any violations of the standard laws governing conjunctions. Since w_5 is among the closest worlds to w_α, however, a world verifying intuitionistic logic but not all of A_1, A_2, and $\neg(A_1 \wedge A_2)$ is *not* closer to w_α than a world—namely w_5—verifying all these sentences. So (14) is false. Similarly, since w_6 is among the closest worlds to w_α, a world verifying intuitionistic logic but not all of B_1, B_2, and $\neg(B_1 \wedge B_2)$ is *not* closer to w_α than a world—namely w_6—verifying all these sentences. So (15) is false. So if (M-ConFac) is to make (1) but not (2) to (6) true—by means of w_5 and w_6—it fails

to make (14) and (15) true. So (M-ConFac) fails when at least two type-3 worlds count as equally close to actuality.

Obviously, we cannot save (M-ConFac) from problems similar to those above by letting at least two different types of impossible worlds candidate for equally close to actuality. So I conclude that (M-ConFac) must give the wrong semantic verdict for some counterlogicals if it is to give the right semantic verdict for (1). Thus (M-ConFac) fails as a general semantics for counterpossibles.

3.1 Remarks on problems

One might observe that all my counterexamples to (M-ConFac) so far have involved reference to intuitionistic logic. But the case against (M-ConFac) is not restricted to such examples. To see this, suppose for simplicity that the uniqueness assumption holds.[13] Consider then any intuitively true counterlogical $A \mathrel{\Box\!\!\rightarrow} C$ that (M-ConFac) makes true. According to (M-ConFac), some logically impossible world w in which A and C are true is then closer to w_α than any logically impossible world in which A is true but C is not. By (Inc), w will verify an instance of a LNC-, CF-, or NCF-inconsistency. As such, we know that (M-ConFac) also must deem at least one of the following counterlogicals *true*:

(G_1) $A \mathrel{\Box\!\!\rightarrow}$ [instance of LNC-inconsistency].

(G_2) $A \mathrel{\Box\!\!\rightarrow}$ [instance of CF-inconsistency].

(G_3) $A \mathrel{\Box\!\!\rightarrow}$ [instance of NCF-inconsistency].

To generate a counterexample to (M-ConFac), we need a logically impossible antecedent A such that $A \mathrel{\Box\!\!\rightarrow} C$ is intuitively true, for some C, but such that (G_1), (G_2), and (G_3) are all intuitively *false* for that antecedent.

Above I specified the relevant impossible antecedent by reference to intuitionistic logic, but other examples are easy to come around. Consider, for instance, the intuitively true counterlogical "If classical logic had not been correct, then some non-classical logic would have been." Suppose (M-ConFac) deems this counterpossible true. Then, by the line of reasoning from above, we know that (M-ConFac) must also

[13]Counterlogicals similar to those that caused problems for (M-ConFac) in the second part of the argument above can easily be generated for the kinds of antecedents that I will consider in this section.

deem at least one of following counterlogicals *true*—bearing in mind that we suppose the uniqueness assumption for simplicity:

(16) (a) If classical logic had not been correct, then [instance of LNC-inconsistency] would have been true.

(b) If classical logic had not been correct, then [instance of CF-inconsistency] would have been true.

(c) If classical logic had not been correct, then [instance of NCF-inconsistency] would have been true.

But all these counterpossibles might well be *false*—say, if intuitionistic logic is correct in the closest impossible world where classical logic is not. So (M-ConFac) cannot be the right semantics for counterpossibles.

We may also focus on logically impossible antecedents that involve reference to certain mildly non-classical logics such as *multi-valued logic* or *supervaluationism*. These logics agree with classical logic on which sets of sentences are satisfiable when each of the relevant sentences receives a designated truth-value: either *true* or *false*. For instance, if $(A \wedge B)$ is assigned the value *true*, then Kleene's 3-valued logic, Łukasiewicz's 3-valued logic, and supervaluationism all agree with classical logic that A and B must also be assigned the value *true*—similar considerations apply to the other standard propositional connectives. Suppose then that (M-ConFac) makes true the intuitively true counterlogical "If 3-valued logic (supervaluationism) were true, then the principle of bivalence would fail." By the line of reasoning from the first part of the argument above, we know that (M-ConFac) must also deem at least one of following counterlogicals *true*:

(17) (a) If 3-valued logic (supervaluationism) were true, then [instance of LNC-inconsistency] would be true.

(b) If 3-valued logic (supervaluationism) were true, then [instance of CF-inconsistency] would be true.

(c) If 3-valued logic (supervaluationism) were true, then [instance of NCF-inconsistency] would be true.

But all these counterlogicals are *false*. Regarding (17a), A and $\neg A$ should not both be designated the value *true*, according to 3-valued logic and supervaluationism. So even if either of these logics were true, a LNC-inconsistency would still not be true. Regarding (17b), if $(A \wedge B)$ is designated the value *true*, then both A and B must be designated the value *true*, according to 3-valued logic and supervaluationism. So even if either of these logics were true, a CF-inconsistency would still not be true. Regarding (17c), if $\neg(A \wedge B)$ is designated the value *true*, A and B should *not* both be designated the value *true*, according to 3-valued logic and supervaluationism. So even if either of these logics were true, a NCF-inconsistency would still not be true. Since (M-ConFac) must deem one of these counterlogicals true, however, it fails as a general semantics for counterpossibles.

4 Alternative world ontologies

I have argued that the extended Lewis-Stalnaker semantics fails when logically impossible worlds are maximal. As far as I can tell, this negative result leaves the proponent of (ConFac) with two options.[14] Either she includes non-maximal or partial worlds in the ontology underlying (ConFac), or she identifies (some) impossible worlds in the ontology underlying (ConFac) with Australian type impossible worlds. Here I will only briefly comment on the first option, but expand on the second. Although I find the second option more interesting than the first, I should not be taken to endorse either. Rather, the current section is mainly exploratory and suggestive of ways that we might develop an extended Lewis-Stalnaker semantics that can avoid the problems that (M-ConFac) had.

4.1 Partial modal space

If we include non-maximal or partial impossible worlds in modal space, we include impossible worlds in which sentences may fail to receive a truth-value—here I assume that at least possible worlds should still be maximal. Call a modal space that includes non-maximal worlds a "partial modal space."

If we want to base an extended Lewis-Stalnaker semantics on a partial modal space, we need firstly determine whether *every* set of sentences should count as a

[14]It will not help the proponent of (M-ConFac) to expand the world-making language to include symbols $\vee$ and $\rightarrow$ that play the same inferential roles as classical disjunction and material implication—and modify (Evaluation) in section 2.1 accordingly. For given the proof of (Inc) in the appendix, it is not hard to see how it generalizes to sentences involving $\vee$ and $\rightarrow$. In turn, it is not hard to verify that arguments similar to those in section 3 involving $\vee$ and $\rightarrow$ can be leveled against (M-ConFac). So for now I set aside the option of enriching the world-making language.

world. If only some sets of sentences count as worlds, we need a principled distinction between sets that do count as worlds and sets that do not. As far as I know, this distinction has not yet been made precise, although the framework in section 4.2 can be seen as an attempt to do so. So for the purpose of this section, I will assume that every set of sentences can count as a world in a partial modal space. This assumption also seems to sit nicely with Nolan 1997 who mentions that if we allow only some impossibilities to count as impossible worlds, then we will find ourselves in

> [...] a distinctly uncomfortable halfway house between those who deny that there are impossible worlds (perhaps the standard position), and on the other hand my position, which maintains that for every impossibility, there is some impossible world where it holds. (Nolan 1997, p. 547.)

Presumably, we would find ourselves in a similar uncomfortable halfway position if we allowed only some sets of sentences to count as impossible worlds.

Given that every set of sentences corresponds to some world in partial modal space, we can define what it means for a sentence to be true and false in a world:

(**Truth***) A sentence A is *true* in a world w iff $A \in w$.
(**Falsity***) A sentence A is *false* in a world w iff $\neg A \in w$.

If $A \notin w$, A is *not* true in w. So if both $A \notin w$ and $\neg A \notin w$, A is neither true nor false in w (it is not true that A is false in w). Alternatively, if $A \notin w$ and $\neg A \notin w$, then A *lacks a truth-value* in w, or A is *indeterminate* in w. For all maximal worlds, A is false in w just in case A is not true in w.

Suppose then that we identify the world-ontology underlying (ConFac) with the class of worlds in partial modal space:[15]

> (**P-ConFac**) $A \;\Box\!\!\rightarrow C$ is true in the actual world w_α iff some world in which A and C are true is closer to w_α than any world in which A is true but C is not.

Since (P-ConFac) is built on a partial world-ontology, (Inc) will no longer follow and the problems discussed in section 3 need no longer arise. And since every set of sentences corresponds to a world, there are worlds in partial modal space that verify

[15]The semantics of (P-ConFac) differs slightly from the semantics of (ConFac) because there are now two ways in which a sentence can fail to be true in a world: either if it is false in the world, or if it is indeterminate in the world.

A but not C, for any two sentences A and C. Accordingly, the ontology underlying (P-ConFac) allows that every counterpossible can be true, and that every counterpossible can be false—perhaps except for $A \boxright A$.

Although I will not engage in a detailed critique of (P-ConFac) here, I will raise a general worry that will serve to motivate the framework in section 4.2. The worry concerns the explanatory power of (P-ConFac). For insofar as every set of sentences corresponds to a world, there is nothing in the construction of worlds—nothing in the underlying world-ontology—that helps reflect the non-trivial semantic, metaphysical, and logical dependencies or relations that obtain between various sentences. Indeed, for any two sentences A and C that stand in some non-trivial semantic, metaphysical, or logical relation to each other, there are worlds in partial modal space in which this relation fails to hold. For there are sets of sentences that contain A but not C, and hence worlds in which A obtains but C does not. In particular, worlds in partial modal space do not allow us to capture any interesting logical relations between sentences. Even though intuitionistic logic and $(A \wedge B)$ had been true, the logical relations that hold intuitionistically between $(A \wedge B)$, A, and B need never be reflected at the level of worlds in partial modal space—there are plenty of partial worlds in which intuitionistic logic and $(A \wedge B)$ are true, but A or B not. But part of the motivation for using worlds in philosophical analyses, I take it, is to reflect such non-trivial logical and inferential relations. Otherwise, it is not obvious which extra explanatory power the appeal to worlds in (P-ConFac) gives us semantically.

Rather, it seems to me, if worlds are to play an explanatory relevant role in a semantics for counterpossibles, they need to have more structure than arbitrary sets of sentences. As Nolan says, "[w]e are often able to say quite exactly what would be the case, logically speaking at least, in the closest impossible world where an actually false logic is true," and it is natural to try and reflect this semantic and logical knowledge explicitly in a world-involving model for counterpossibles (Nolan 1997, p. 545). For instance, in the closest impossible world where intuitionistic logic and $(A \wedge B)$ are true, there should be a fact about the construction of worlds that guarantees that A and B are also true in that world. If not, we may as well bypass talk of worlds and explicate the semantics for counterpossibles without appealing to these theoretical entities.

So if we can find a better way to account for counterpossibles, I think we should. Insofar as we need appeal to partial worlds to avoid the negative results of section

3, proponents of the extended Lewis-Stalnaker semantics owe us an account of the nature of these worlds. In light of the above, I take it, such an account should impose substantial constraints on which sets of sentences can count as worlds. In the next section, I will sketch a framework that goes some way towards meeting this request.

4.2 Stratified modal space

In the previous section I motivated the thought that worlds—whether possible or impossible, and whether maximal or non-maximal—should have substantially more structure than what arbitrary sets of sentences offer. In this section, I will show how Australian type impossible worlds might give us this additional structure and how an extended Lewis-Stalnaker semantics based on such worlds can avoid the problems discussed in section 3.

Generally speaking, Australian type impossible worlds have more structure than American type impossible worlds. For whereas the latter correspond to arbitrary—or perhaps arbitrary maximal—sets of sentences, the former correspond to sets of sentences that are closed under logical consequence in some non-classical logic. Since we will want to quantify over more than one non-classical logic, it is best not to specify the nature of the relevant Australian type impossible worlds by appeal to just one specific non-classical logic. Instead I will adopt the following liberal definition of what it means to be a world:

> (**World**) A set Γ of sentences (in the world-making language) is a world w iff Γ is closed under logical consequence in logic $\mathcal{L}_i$.

Whereas logically possible worlds correspond to maximal sets of sentences that are closed under classical logical consequence—except for the "explosion" world that corresponds to the maximal, inconsistent set of all sentences—logically impossible worlds now correspond to sets of sentences that are closed under some non-classical notion of logical consequence. Henceforth, whenever I use the word "world" without qualification, I will mean the types of (Australian) worlds that are defined by (World).

Given this conception of worlds, we can now construct an ontology that consists of a spectrum of modal spaces $W_{\mathcal{L}_1}, \ldots, W_{\mathcal{L}_n}$, where:

> (**Strat**) $W_{\mathcal{L}_i}$ = the class of sets of sentences (in the world-making language) that are closed under logical consequence in logic $\mathcal{L}_i$.

To construct the space of logically possible worlds, we take the class of all maximal sets of sentences and close each of them under logical consequence in classical logic.[16]

Since impossible worlds now will be typed according to various non-classical logics, we can also refer to them as worlds that are possible with respect to some non-classical logic—bearing in mind the assumption that the only *genuine* logically possible worlds are the ones that respect classical logic. To construct the space of intuitionistically possible worlds, we take the class of all sets of sentences and close each of them under logical consequence in intuitonistic logic. And to construct the space of paraconsistently possible worlds, we take the class of all sets of sentences and close each of them under logical consequence in paraconsistent logic. Call the modal space that contains all the spaces of worlds that are possible relative to each logic $\mathcal{L}_i$ "stratified modal space".[17]

Given this construction of stratified modal space, we need to determine what it means for a sentence to be true in a world. Insofar as we want to retain a set-theoretical account of truth and falsity in a world, we need to change the basic picture from section 2.1 to accommodate the kinds of truth-values that are licensed by various non-classical logics. In particular, it seems, we cannot easily retain (Truth) and (Falsity) for spaces of worlds that model non-classical logics such as fuzzy logic in which there are continuum many truth-values.

To overcome the limitations of the simple set-theoretical account of truth-in-a-world, we may use elements from fuzzy set theory and allow sentences to have degrees of membership in the sets of sentences that correspond to worlds.[18] Let $\mu_w(A)$ represent the degree x—for some x in the real interval $[0, 1]$—to which A belongs to the world w. We can then define:

---

[16]Since classical logic has an explosive consequence relation, every classically inconsistent set will be identical to the set that contains every sentence in the world-making language. This "explosive" set of sentences will not correspond to any logically possible world, and henceforth we should assume that it is excluded from the space of logically possible worlds.

[17]For now I leave it an open question whether we want to quantify over only existing logics, or also over conceivable logics.

[18]In claiming that we can fruitfully utilize certain aspects of fuzzy set theory, I only want to claim that we can utilize the idea of degrees of set membership but not that we have to endorse the whole fuzzy set-theoretical package.

(**D-Truth**)	$\mu_w(A) = 1$	iff	A is *true* in w.
(**D-Falsity**)	$\mu_w(\neg A) = 1$	iff	A is *false* in w.
(**De-Truth**)	$\mu_w(A) = x$	iff	A is *true in w to degree x*, for $x \in (0, 1)$.
(**De-Falsity**)	$\mu_w(\neg A) = x$	iff	A is *false in w to degree x*, for $x \in (0, 1)$.

If $\mu_w(A) = 0$, A is *not true* in w. Since we need to countenance various non-classical logics in the framework, "not true" is detached from "falsity" or "truth of negation". If $\mu_w(A) = 1$, A is wholly included in w, and if $\mu_w(A) = 0$, A is wholly excluded from w.

For classical logics and their close 2-valued cousins, "not true" and "falsity" or "truth of negation" will collapse: $\mu_w(\neg A) = 1$ just in case $\mu_w(A) = 0$. In such cases, the set-theory will be indistinguishable from classical set-theory, and (De-Truth) and (De-Falsity) will not play a role. For logics that acknowledge only *truth* and *falsity* as designated truth-values but nonetheless operate with truth-value gaps—as in the model in section 4.1—we need spaces of worlds such that $\mu_w(A) = 0$ and $\mu_w(\neg A) = 0$. Interpreted, there will be worlds in the relevant spaces in which A is neither true nor false (not true that $\neg A$). For logics with more than two designated truth-values, there are various options. For a standard 3-valued logic, we might let the third designated truth-value that a sentence can take in a world be given by $\mu_w(A) = 0.5$. For logics with finite or continuum many truth-values, we can use the full resources of the degree-theoretical account of what it means for a sentence to be true in a world.

Rather than appealing to fuzzy set theory, one may also consider taking "truth-in-a-world" as a primitive notion in the framework and define an evaluation function Val from pairs of worlds and sentences into truth-values in the interval $[0, 1]$:

(**D-Truth***)	$Val(w, A) = 1$	iff	A is *true* in w.
(**D-Falsity***)	$Val(w, \neg A) = 1$	iff	A is *false* in w.
(**De-Truth***)	$Val(w, A) = x$	iff	A is *true in w to degree x*, for $x \in (0, 1)$.
(**De-Falsity***)	$Val(w, \neg A) = x$	iff	A is *false in w to degree x*, for $x \in (0, 1)$.

Depending on which spaces of worlds we focus on, the properties of Val may stand in further relations to each other. For instance, in spaces where negation behaves like classical negation, $Val(w, \neg A) = 1$ just in case $Val(w, A) = 0$.

Whether or not we opt for a set-theoretical or primitive representation of truth-in-a-world—or yet some other representation—the evaluation of complex sentences in worlds will depend on the semantics for the logic that governs the worlds in question.

As such, there will most often be additional constraints on the evaluation of sentences in worlds when we focus on specific spaces in stratified modal space.

Let now "(S-ConFac)" refer to a version of the extended Lewis-Stalnaker semantics that is formulated just like (P-ConFac) but undergirded by stratified modal space.[19] Since worlds are located within specific spaces or spheres of worlds in stratified modal space, closeness intuitions pertain no longer just to worlds but also to spaces of worlds. Although relative closeness between worlds and between modal spaces will largely be determined intuitively and contextually, we can be a bit more specific than that.

First, we can identify the modal space $W_{\mathcal{L}_1}$ in the spectrum $W_{\mathcal{L}_1}, \ldots, W_{\mathcal{L}_n}$ with the space of logically possible worlds in which the actual world w_α is a member. We can then assume a version of Nolan's:[20]

> (**Strangeness of Impossibility Condition**) For any logically possible world w in $W_{\mathcal{L}_1}$ and any (classically) logically impossible world $w^\star$ in $W_{\mathcal{L}_n}$, for $n > 1$, w is closer to w_α than $w^\star$.

On the face of it, this condition seems reasonable. It captures the idea that possible worlds are always closer to the actual world than any impossible world. As Nolan puts it, "[t]he heavens will fall before (correct) logic fails us." (Nolan 1997, p. 550.)

Second, we can impose the following condition on relative closeness between modal spaces:

> (**Relative Closeness Condition**) For any counterfactual whose antecedent presupposes that some logic $\mathcal{L}_i$ is correct (true, adequate), a world in modal space $W_{\mathcal{L}_i}$ is closer to the actual world than any world in modal space $W_{\mathcal{L}_j}$, where $W_{\mathcal{L}_i} \neq W_{\mathcal{L}_j}$, and where $i \geq 1$ and $j > 1$.

This condition too seems reasonable. It captures the idea that if some logic $\mathcal{L}_i$ had indeed been correct, then regardless of what else might have been the case, the laws of $\mathcal{L}_i$ would have been the case. For instance, if intuitionistic logic had been correct, then regardless of what else might have been the case, the world would have been a place where the laws of intuitionistic logic hold.

[19] If we were to allow evaluation of counterfactuals across arbitrary worlds in stratified modal space, we would have to allow for more a complex—perhaps even degree-theoretical—semantics for counterfactuals. As long as we assume that the actual world is a world in which classical logic holds, however, we can forget about these other truth-values in explicating (ConFac).

[20] Cf. Nolan 1997, p. 550.

Finally, notice that the current framework validates:[21]

(**Entailment Principle**) If C is a logical consequence of A in logic $\mathcal{L}_i$, then $A \mathrel{\Box\!\!\to} C$ is true for any counterfactual whose antecedent presupposes that $\mathcal{L}_i$ is correct (true, adequate).

To see that (Entailment Principle) holds, suppose that C is a logical consequence of A in logic $\mathcal{L}_i$. Consider then $A \mathrel{\Box\!\!\to} C$, where A presupposes that $\mathcal{L}_i$ is correct. To evaluate $A \mathrel{\Box\!\!\to} C$, we know by (Relative Closeness Condition) that a world in $W_{\mathcal{L}_i}$ is closer to w_α than any world in some distinct modal space $W_{\mathcal{L}_j}$. By construction, every world in $W_{\mathcal{L}_i}$—whether close to w_α or not—that verifies A also verifies C. So $A \mathrel{\Box\!\!\to} C$ is true, according to (S-ConFac). Hence (Entailment Principle) holds.

Given the features above, it is easy to see why the problems that I raised for (M-ConFac) no longer trouble (S-ConFac). Consider again (1):

(1) If intuitionistic logic were correct and $(A \wedge B)$ true, then A and B would also be true (for any A and B).

As it should, (S-ConFac) renders (1) true. Since the antecedent in (1) presupposes that intuitionistic logic is correct, we know by (Relative Closeness Condition) that some world in intuitionistic modal space $W_{\mathcal{L}_{IL}}$ is closer to w_α than any other world in some other non-classical space. And since each world in $W_{\mathcal{L}_{IL}}$ is closed under logical consequence in intuitionistic logic, we know that any world in $W_{\mathcal{L}_{IL}}$—whether close to w_α or not—that verifies $(A \wedge B)$ also verifies A and B. So (1) is true, according to (S-ConFac). Alternatively, since A and B are logical consequences of $(A \wedge B)$ in intuitionistic logic, we know by (Entailment Principle) that (S-ConFac) will deem (1) true.

In contrast to (M-ConFac), however, (S-ConFac) can render (1) true without thereby also erroneously rendering true the false counterpossibles (2) to (6). To see this, consider again (2):

(2) If intuitionistic logic were correct and $(A \wedge B)$ true, then C and $\neg C$ would both be true (for some C).

[21] (Entailment Principle) is a version of a similar principle in the standard Lewis-Stalnaker semantics that allows us to infer $A \mathrel{\Box\!\!\to} C$ whenever C is a (classical) logical consequence of A.

For (2) to be true, according to (S-ConFac), we would need a world in intuitionistic modal space $W_{\mathcal{L}_{IL}}$ that verifies both C and $\neg C$. Since each world in $W_{\mathcal{L}_{IL}}$ is closed under logical consequence in intuitionistic logic, a world in $W_{\mathcal{L}_{IL}}$ that verifies a contradiction will verify *all* other sentences as well. For contradictions entail everything in intuitionistic logic. But such an "explosion" world—a world in which every sentence is true—seems very far away from w_α. Indeed, as Nolan says, "it seems to be one of the most absurd situations conceivable." (Nolan 1997, p. 544.) If so, any world in $W_{\mathcal{L}_{IL}}$ that verifies at *most* one member of a contradictory pair of sentences is closer to w_α than the world in $W_{\mathcal{L}_{IL}}$ that verifies a contradiction and hence everything else. So (2) comes out false, according to (S-ConFac), as it should.[22] Since we can give a similar account for the falsity of the other troublesome counterpossibles in section 3, (S-ConFac) is not troubled by the problems that (M-ConFac) had.

Although (S-ConFac) can avoid the problems of section 3, people have complained that an extended Lewis-Stalnaker semantics based on Australian type impossible worlds cannot work. For instance,

> Daniel Nolan points out that a uniform weakening of the consequence relation is a bad idea, because no weakening can handle every impossibility that we might want to reason about. Here is a way to argue that very general point in terms of counterpossibles. Let L be the preferred/correct but weakened logic. Then, absurdly, the following counterpossible is not false:
>
> > [(18)] If L were not the correct logic, then all and only L-theorems would be valid.
>
> The conditional would not be treated as false, because ex hypothesi every world would be an L-world. Hence, there would be no worlds where the consequent is false, and so, no closest antecedent worlds where the consequent is false. Of course, L is an arbitrary logic. So any uniform weakening of the consequence relation would fail in the same way. By itself the strategy fails to capture all reasonable intuitions about the truth-values of counterpossibles. (Brogaard & Salerno 2013, p. 651.)

While this objection has force against a version of the extended Lewis-Stalanker semantics that is based on a *single* space of Australian type impossible worlds, it does not have force against (S-ConFac). For the purpose of evaluating the counterlogical

[22] Alternatively, we can make (2) false, according to (S-ConFac), by stipulating that the explosion world is not in $W_{\mathcal{L}_{IL}}$.

(18), we should go to a space of worlds that is not governed by the logic L in question.[23] And in such a space, the consequent in (18) need not be true, and hence (18) need not be true, according to (S-ConFac).

One may also complain that we can reason about impossibilities in the absence of *any* logic, and hence that we can entertain counterpossibles that do not respect *any* logical constraints.[24] If so, it seems that we need logically impossible worlds that are not governed by any logic to evaluate certain counterpossibles, and hence that we need logically impossible worlds that are not of the Australian type. Given that all impossible worlds in the ontology underlying (S-ConFac) are of the Australian type, (S-ConFac) will not be able to handle all counterpossibles that we might care about.

In reply, notice that nothing I have said prevents us from isolating a space of worlds that are closed under logical consequence in some ultra-weak logic $\mathcal{L}_X$ with no—or hardly any—principles governing logical consequence. For all practical intents and purposes, we can think of the corresponding modal space $W_{\mathcal{L}_X}$ as a space in which every set of sentences corresponds to a world. Insofar as every—or almost every—set of sentences corresponds to a world in $W_{\mathcal{L}_X}$, the semantic behavior of sentences in worlds in this space is effectively completely unconstrained. As such, everything goes in $W_{\mathcal{L}_X}$, both semantically and logically speaking. To evaluate counterpossibles that do not respect any logical constraints, we can hence focus on worlds in $W_{\mathcal{L}_X}$ and avoid the objection above.[25]

In some sense, of course, the reply above still assumes that counterpossible reasoning takes place in some logic. But since the logic $\mathcal{L}_X$ can be arbitrarily weak and formally unconstrained, there is no harm in making this assumption for the purpose of giving a semantics for counterpossibles—although there might be for the purpose of giving an account of impossible reasoning. Alternatively, we may expand stratified modal space with a particular sphere of American type impossible worlds, including

[23]Notice that the antecedent in (18) does not assume that the logic L is correct (true, adequate). As such, (Relative Closeness Condition) does not apply in this context.

[24]Nolan raises an objection along those lines (Nolan 1997, pp. 547-548).

[25]Notice also that the specification of $W_{\mathcal{L}_X}$ allows us to validate extensionally the following principle that Nolan 1997 tentatively subscribes to:

> (**Comprehension Principle**) For any set Δ of sentences that cannot possibly be jointly true, there is an impossible world in which each sentence in Δ is true.

For given that there are no—or hardly any—principles governing logical consequence in $\mathcal{L}_X$, any set of sentences—or almost any set of sentences—will count as a world in $W_{\mathcal{L}_X}$.

both maximal and non-maximal ones. To evaluate counterpossibles that do not respect any logical constraints, we could then appeal to worlds in this space of American type impossible worlds to avoid the objection above.

If we are liberal enough about the logics that we use to construct the different modal spaces, the stratified ontology underlying (S-ConFac) can hence be as liberal as ontologies based on arbitrary maximal and non-maximal American type worlds. So (S-ConFac) will be as tolerant to impossibilities as (M-ConFac) and (P-ConFac), and be able to capture as many fine-grained semantic distinctions as these two alternative accounts. But (S-ConFac) also seems to fare better than both (M-ConFac) and (P-ConFac): better than (M-ConFac) because it avoids the problems in section 3, and better than (P-ConFac) because it has a non-trivial formal structure that validates interesting principles.

Yet, let me stress that I have not attempted to give a full-scale defense—nor a full-scale exposition—of a semantics for counterpossibles. I have merely discussed two ways in which we may develop an extended Lewis-Stalnaker semantics that can avoid the problems of (M-ConFac), and indicated my reasons for preferring (S-ConFac) over (P-ConFac). Whether the construction of stratified modal space survives closer scrutiny, or whether it needs adjusting in crucial respects, is a topic for another paper. For now (S-ConFac) is on the table as a new rival to the existing semantic frameworks for counterpossibles.

5 Conclusion

If we identify logically impossible worlds with maximal, logically inconsistent sets of sentences, the extended Lewis-Stalnaker semantics for counterfactuals fails: it gives the wrong semantic verdicts for many counterlogicals. Arguably, the standard version of the extended semantics presupposes not only that impossible worlds are American type impossible worlds, but also that they are maximal. If so, my arguments show that the standard way of making the extended Lewis-Stalnaker semantics precise fails.

To avoid the negative result, a proponent of the extended Lewis-Stalnaker semantics may include non-maximal American type impossible worlds in the underlying world-ontology. By doing so, she may avoid the problems from section 3. I argued that there are reasons to impose substantial logical constraints on which (partial) sets of sentences should count as worlds, and went on to investigate a stratified world-ontology consisting of Australian type impossible worlds most of which obey such constraints. I argued that an extended Lewis-Stalnaker semantics built on a stratified modal space can avoid the problems from section 3, but also that it remains as tolerant to impossibilities as ontologies consisting of maximal and non-maximal American type impossible worlds. As such, the proposal can avoid certain problems that other existing semantics for counterpossibles have, while at the same time make as many fine-grained semantic distinctions as the existing semantics.

Although I have not aimed to give a detailed defense of the extended Lewis-Stalnaker semantics, I have aimed to give detailed recommendations for proponents of the semantics. For insofar as a successful Lewis-Stalnaker semantics for counterpossibles cannot be built on an ontology consisting solely of maximal worlds, my arguments recommend that it be built upon either partial American type impossible worlds or Australian type impossible worlds. I have outlined one such positive account of counterpossibles, but a full-scale exposition must wait for future work.

6 Appendix

In this appendix, I give the proofs of (Satisfiability) and (Inc).

Proof of (Satisfiability)

Let Γ be any set of sentences that satisfies (i) and (ii) in (Satisfiability). We want to show that there is a propositional evaluation function ν that makes each sentence A in Γ true. To this end, we stipulate an interpretation $\mathcal{I}$ such that for all atomic A:

$\mathcal{I}(A) = T$ iff $A \in \Gamma$.

$\mathcal{I}(A) = F$ iff $A \notin \Gamma$.

This is a possible stipulation because $\mathcal{I}$ cannot assign both T and F to any atomic A. We then need to show that every sentence in Γ is true under this interpretation. I do this by induction on the length of a sentence, where the length of a sentence is given by the number of symbols it contains:

Base case: Assume for atomic A that $A \in \Gamma$. We want to show that $\nu(A) = T$. We get the result immediately. By definition of $\mathcal{I}$, $A \in \Gamma$ iff $\mathcal{I}(A) = T$. By ($\nu\mathcal{I}$), $\mathcal{I}(A) = T$ iff $\nu(A) = T$. So $A \in \Gamma$ iff $\nu(A) = T$. So $\nu(A) = T$.

Inductive step: Assume for the induction hypothesis that every sentence in Γ that is shorter than $\neg A$ and $(A \wedge B)$ is true under the evaluation ν based on $\mathcal{I}$. We want to show that if $\neg A \in \Gamma$, then $\nu(\neg A) = T$, and if $(A \wedge B) \in \Gamma$, then $\nu(A \wedge B) = T$. There are two cases to consider:

Case 1: Assume $\neg A \in \Gamma$. By (i) in (Satisfiability), $\neg A \in \Gamma$ iff $A \notin \Gamma$. By induction hypothesis, $A \notin \Gamma$ iff $\nu(A) = F$. By ($\nu\neg$) in (Evaluation), $\nu(A) = F$ iff $\nu(\neg A) = T$. So $\nu(\neg A) = T$.

Case 2: Assume $(A \wedge B) \in \Gamma$. By (ii) in (Satisfiability), $(A \wedge B) \in \Gamma$ iff $A \in \Gamma$ and $B \in \Gamma$. By induction hypothesis, $A \in \Gamma$ and $B \in \Gamma$ iff $\nu(A) = T$ and $\nu(B) = T$. By ($\nu\wedge$) in (Evaluation), $\nu(A) = T$ and $\nu(B) = T$ iff $\nu(A \wedge B) = T$. So $\nu(A \wedge B) = T$. $\square$

Proof of (Inc)

Let Δ be any maximal, logically inconsistent set of sentences. By (Satisfiability), Δ will fail to satisfy either (i) or (ii) and hence contain at least one of the following inconsistent pairs or triples of sentences:

Case 1: Δ may be inconsistent because it fails to satisfy (i) of (Satisfiability), in which case Δ contains an inconsistency of the form $\{A, \neg A\}$. That is, Δ contains an instance of a LNC-inconsistency.

Case 2: Δ may be inconsistent because it fails to satisfy (ii) of (Satisfiability), in which case Δ contains either an inconsistency of the form $\{\neg A, (A \wedge B)\}$, $\{\neg B, (A \wedge B)\}$, or $\{\neg A, \neg B, (A \wedge B)\}$, or an inconsistency of the form $\{A, B, \neg(A \wedge B)\}$. That is, Δ contains either an instance of a CF-inconsistency or an instance of a NCF-inconsistency. □

References

Adams, R. (1974). Theories of Actuality. In *Nous*, 8: pp. 211-231.

Berto, F. (2009). Impossible Worlds. In *The Stanford Encyclopedia of Philosophy* (Winter 2003 Edition). Edward N. Zalta (ed.). URL = `http://plato.stanford.edu/entries/impossible-worlds/`.

Brogaard, B. & Salerno, J. (2013). Remarks on Counterpossibles. In *Synthese*, 190, pp. 639-660.

Carnap, R. (1947). *Meaning and Necessity, a Study in Semantics and Modal Logic*. Chicago: University of Chicago Press.

Chalmers, D. J. (2011). The Nature of Epistemic Space. In A. Egan, & B. Weatherson (Eds.), *Epistemic Modality*. New York: Oxford University Press, pp. 60-107.

Goodman, J. (2004). An Extended Lewis/Stalnaker Semantics and the New Problem of Counterpossibles. In *Philosophical Papers*, 33, pp. 35-66.

Hintikka, J. (1969). *Models for Modalities*. Dordrecht: D. Reidel Publishing Company.

Jeffrey, R. (1983). *The Logic of Decision*. Chicago: University of Chicago Press.

Krakauer, B. (2012). What are impossible worlds? In *Philosophical Studies*, online first, pp. 1-19.

Lewis, D. (1973). Counterfactuals and Comparative Possibility. In *Journal of Philosophical Logic*, 2, pp. 418-446.

Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell Publishers.

Mares, E. (1997). Who's Afraid of Impossible Worlds? In *Notre Dame Journal of Formal Logic*, 38, pp. 516-526.

Nolan, D. (1997). Impossible Worlds: A Modest Approach. In *Notre Dame Journal for Formal Logic* 38, pp. 325-527.

Stalnaker, R. (1968). A Theory of Conditionals. In *Studies in Logical Theory, American Philosophical Quarterly* Monograph Series, 2. Oxford: Blackwell, pp. 98-112.

Vander Laan, D. (2004). Counterpossibles and Similarity. In F. Jackson, & G. Priest (Eds.), *Lewisian Themes*. Oxford: Oxford University Press, pp. 258-276.

Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell.