

## Article

## Ambifictional Counterfactuals

Andrew D. Bassford 

Philosophy Department, University of Texas at Austin, Austin, TX 78713, USA; a.d.bassford@utexas.edu

**Abstract:** In this paper, I argue that David Lewis's possible world semantics for counterfactual discourse and for fictional discourse are apparently inconsistent and in need of revision. The problem emerges for Lewis's account once one considers how to evaluate *ambifictional counterfactuals*. Since this is likely not a concept familiar to most, and since it does not appear that the problem has been previously recognized in the critical literature, I will begin by rehearsing Lewis's possible worlds' semantics for counterfactuals and fiction. Then I will explicate the distinction between intrafictional, extrafictional, interfictional, and ambifictional propositions. Next, I will state what an ambifictional counterfactual proposition is, and explain why this kind of discourse confounds Lewis's system. I will conclude, finally, with a brief discussion of how the Lewisian would be best advised to resolve the paradox.

**Keywords:** possible world semantics; possible world theories of counterfactuals; truth in fiction; fictional characters; David Lewis

## 1. Counterfactual Propositions

A counterfactual proposition is a subjunctive conditional proposition with an assumed false antecedent. Here is an example of a counterfactual proposition:

α Had I been more successful at dating during my undergraduate career, I would not have spent so much time reading David Lewis and watching Star Trek ([1,2]).

There are two types of counterfactual proposition: *would-have* counterfactuals and *might-have* ones. The former are symbolized as ' $\varphi \Box \rightarrow \psi$ ', and read imperfectly as "Had  $\varphi$  obtained,  $\psi$  would have obtained". And the latter are symbolized as ' $\varphi \Diamond \rightarrow \psi$ ', and read imperfectly as "Had  $\varphi$  obtained,  $\psi$  might have obtained". So, understood, (α) is a would-have counterfactual and would be approximately symbolized as 'I am more successful at dating  $\Box \rightarrow \sim(I \text{ read David Lewis} \wedge I \text{ watch Star Trek})$ '.

Under what conditions is a claim like (α) true? Lewis proposed possible worlds' semantics for evaluating the class. According to Lewis ([3–8]; cf. [9,10]):

- $v(<\varphi \Box \rightarrow \psi, w>) = 1$  iff among all the worlds closest to  $w$  at which  $v(\varphi) = 1$ , all of them are such that  $v(\psi) = 1$  too;  $v = 0$  otherwise.
- ◇  $v(<\varphi \Diamond \rightarrow \psi, w>) = 1$  iff among all the worlds closest to  $w$  at which  $v(\varphi) = 1$ , some of them are such that  $v(\psi) = 1$  too;  $v = 0$  otherwise.

More informally, Lewis tells us to evaluate a proposition like (α) as follows (cf. [11–14]). We start by fixing the initial point of evaluation to—in most ordinary cases—the actual world,  $w@$ . Next, we examine all of the possible worlds in the modal galaxy and deselect from our model of logical space any worlds at which I am not more successful at dating during my undergraduate career. Second, among those remaining worlds, we then deselect from the set of worlds any which are not *closest* to  $w@$ . 'Closeness' here is understood by Lewis in terms of similarity, such that  $w_1$  is closest to  $w_2$  just in case  $w_1$  is more similar to  $w_2$  than any other world (meeting the specified conditions). Two worlds might be similar to one another in terms of law (logical, metaphysical, physical, moral, etc.) or in terms of fact. Lewis advises us to use a weighted metric to calculate the strength of similarity among worlds. He thought that nomic similarity counts for more than factual similarity;



Citation: Bassford, A.D.

Ambifictional Counterfactuals.

Philosophies 2023, 8, 108. <https://doi.org/10.3390/philosophies8060108>

doi.org/10.3390/philosophies8060108

8060108

Academic Editor: Frederick Kroon

Received: 16 September 2023

Revised: 7 November 2023

Accepted: 9 November 2023

Published: 12 November 2023



Copyright: © 2023 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

creativecommons.org/licenses/by/

4.0/).

but that next in order of importance is for two worlds to contain a maximally similar timeline leading up to when  $\phi$  should first obtain. So, for example, if we are evaluating  $(\alpha)$  relative to  $w@$ , and we have already deselected from the model any worlds at which the antecedent is false, Lewis tells us to next deselect all *miracle* worlds—i.e., worlds which do not behave in accordance with our laws, and then deselect all *backtracking* worlds—i.e., worlds which do not share the same sequence of events, properly ordered in time, leading up to the antecedent's first obtaining. The first will exclude worlds where gravity fails or where  $3^2 = 1$  (if such worlds exist) or where the principle of sufficient causation does not hold. And the second will exclude all worlds where I was born a Klingon or where I am as handsome as Captain James T. Kirk or where I am as charismatic as Captain Jean-Luc Picard. The remaining world (or worlds) will be the closest to  $w@$ . Finally, we evaluate the consequent at those worlds; if they are all worlds at which the consequent of  $(\alpha)$  is true, then  $(\alpha)$  is true; if not, then not.

## 2. Intrafictional Propositions

A fictional proposition is a proposition that refers to fictional entities. There seem to be four kinds of fictional proposition: intrafictionals, extrafictionals, interfictionals, and ambifictionals. An *intrafictional* is a fictional proposition that ascribes to a fictional entity some property which it is assumed to possess within its native fictional context. Here is an example of an intrafictional proposition:

$\beta$  Lieutenant Commander Data, the android, was created by Dr. Noonian Soong.

Under what conditions is a claim like  $(\beta)$  true? Lewis likewise proposed a possible worlds' semantic for evaluating this class of statement. According to Lewis ([15]), propositions like  $(\beta)$  are not purely extensional, as their surface grammar might suggest, but are rather intensional and need to be evaluated as though they came equipped with an implicit 'According to the fiction of. . .' modal operator (cf. [16–21]). So, in the language of our modal propositional calculus, intrafictionals like  $(\beta)$  should not be symbolized simply as ' $\phi$ ' but rather as ' $f(\phi)$ ', where ' $f$ ' is our new unary intensional operator. Moreover, since there are many different fictions, the operator will always need to be indexed to some fictional context,  $C$ . With this new operator in hand, Lewis then proposed the following:

$f_C \quad v(<f_C(\phi), w>) = 1$  iff among all the worlds closest to  $w$  at which the events detailed in  $C$  are told as known fact, and all of them are such that  $v(\phi) = 1$ ;  $v = 0$  otherwise.

So, understood,  $(\beta)$  has the form ' $f_{ST}(\text{Data was created by Dr. Soong})$ ', where  $ST$  denotes the Star Trek fiction; and it is true just in case, according to the nearest worlds where Star Trek is told as known fact, Data is Soong's created android.

Two points about this proposed semantic should be noted before moving on to other varieties of fictional discourse. The first point is that the concept of modal nearness (closeness, similarity) appears again, just as it did in  $(\Box \rightarrow)$  and  $(\Diamond \rightarrow)$ . This is not surprising; as Lewis [15] (pp.42) notes, "[r]easoning about truth in fiction is very like counterfactual reasoning," and so we should expect their respective semantics to represent that corresponding similarity too. In "Truth in Fiction," Lewis does not state explicitly how, or whether, fictional similarity among worlds is like or unlike counterfactual similarity. Nonetheless, given that his weighted criterion for counterfactual similarity came later in his career [5,7] (in 1979 and 1986 respectively), I believe we would be charitable to read it back into his 1978 work and presume that all that has been said about counterfactual similarity holds true also of fictional similarity. The reason Lewis cites for including this criterion in his semantic is that, even though fictions often fail to disclose at length to what extent their worlds, or "universes", correspond to ours in terms of nomic and factual similarity, it would be gratuitous for an interpreter of a fiction to suppose that many of our fundamental laws and facts do not similarly obtain within our fictions. In the case at hand, Star Trek never specifies whether, e.g., the law of conservation of mass/energy obtains; nonetheless, it would be pernicious to include in our evaluation of Data being an android any worlds where the law does not obtain, unless explicitly stated otherwise. Consider, too, that many

literary theorists presume it is possible to offer psychoanalytic interpretations of literary works (cf. [22]: §5), but the feasibility of this project depends on the assumption that literary universes share a great many of our psychological laws, which need not be stated explicitly in the story.

The second point is this idea of a fiction “being told as known fact” in the target worlds. On this point, I need not say much; Lewis himself is already clear on why he thinks this criterion is necessary (see [15] pp.43–45). The present argumentative purpose demands only that I say this: it is generally assumed that if something is indeed known fact, then it must also be true. So, a weaker version of  $(f_C)$  is as follows:

$f_C^* \quad v(<f_C(\varphi), w>) = 1$  iff among all the worlds closest to  $w$  at which  $v(C) = 1$ , all of them are such that  $v(\varphi) = 1$  too;  $v = 0$  otherwise.

Moving forward, I will therefore speak of a tension between  $(\Box \rightarrow / \Diamond \rightarrow)$  and  $(f_C^*)$ . But since  $(f_C^*)$  is a consequence of  $(f_C)$ , what holds here will hold also between  $(\Box \rightarrow / \Diamond \rightarrow)$  and  $(f_C)$ .

### 3. Extrafictional Propositions

An *extrafictional* is a fictional proposition that refers to a fictional entity qua fictitious—i.e., as a human artifact. Here is an example of an extrafictional proposition:

$\gamma$  Lieutenant Commander Data, the character, was created by Gene Roddenberry.

Under what conditions is a proposition like  $(\gamma)$  true? On this point, Lewis [15] (pp.38) says only that “I shall have nothing to say here about the proper treatment of these sentences. If the Meinongian can handle them with no special dodges, that is an advantage of his approach over mine” (cf. [23]). But this is not a very satisfying response, especially since Lewis does not then go on to specify what those special dodges might be. So, for present dialectical purposes, I will need to go off script and reason independently about how I think the Lewisian would do best to semantically model fictional propositions of this sort.

I propose the Lewisian accommodate extrafictionals into their theory by making two moves. The first move is to suppose that the actual domain includes not only concrete objects, but also abstract objects, and that some of the latter may be artificial (cf. [24]). (Note that Lewis himself already made provisions in his fundamental ontology for abstracta, like universals—see [25].) And the second move is to suppose that fictional names are referentially ambiguous, in that they may refer either to concrete possibilia, on the one hand, or to actual abstracta, on the other. Intrafictionals denote the former; extrafictionals denote the latter. So, understood, we can say that a name like ‘Data’—as in  $(\beta)$  and  $(\gamma)$ —is referentially ambiguous too. It may refer to an android or to certain abstract intellectual property. We might disambiguate this within our language by adding a subscript to each name: ‘Data<sub>f</sub>’ and ‘Data<sub>@</sub>’.  $(\beta)$  references Data<sub>f</sub>, whereas  $(\gamma)$  references Data<sub>@</sub>.

I believe this proposal can be motivated by considering some important remarks Lewis makes in his *On the Plurality of Worlds*. There, Lewis considers the possibility of interworld causation and rejects it. The reason he rejects it is because Lewis supported a counterfactual theory of causation, and he very strongly rejects the possibility of counterfactual dependency relations obtaining among objects/events from different worlds. (For present purposes, we need not follow him to the letter on this point; it is enough if many causal relations, such as creation, merely entail counterfactual dependency in order for his point to stick.) Lewis’s point was just that if we allow for something in one world to counterfactually depend on something in another world, our semantic quickly breaks down. As Lewis [7] (pp. 78–79) puts it, suppose that “[e]vent  $C$  occurs at world  $W_C$ , event  $E$  occurs at world  $[W]_E$ , they are distinct events, and if  $C$  had not occurred,  $E$  would not have occurred either. This counterfactual is supposed to hold—where? It means that at the closest worlds to—where?—at which  $C$  does not occur,  $E$  does not occur—where?—either”. Whether we understand the worlds closest relative to  $W_C$  or  $W_E$ , we reach a counterintuitive result. It is of no help to understand the counterfactual as saying, e.g., “at the closest worlds to  $W_C$  at which  $C$  does not occur at  $W_C$ ,  $E$  does not occur in  $[W]_E$ ”, since that is like saying

“in Auckland it rains in Melbourne, but in Wellington it doesn’t rain in Melbourne,” for which no sense can be made. Nor do any alternative evaluation schemes seem plausible. Hence, our best option is to deny interworld causation so that we will have denied a class of relations that would otherwise imply interworld counterfactual dependence. (More on this later.)

Now, if the Lewisian does not accept my proposal for making sense of extrafictionals, then they face a reductio. Take  $(\gamma)$ , which names ‘Data’. If this Data is not an abstract object, then it must be the same ‘Data’ as  $\text{Data}_f$ , since the statement is apparently true and there are no androids in  $w@$ . But if so, then  $(\gamma)$  states that Roddenberry created an object from a different world, which is a kind of causal relation, which would imply counterfactual dependency between events from two different worlds, which would cause our semantic to break down in just the way Lewis thought. Hence, we should suppose that ‘Data’ in  $(\gamma)$  refers to an actual abstract object from  $w@$ —viz.  $\text{Data}_@$ .

#### 4. Interfictional Propositions

An *interfictional* is a fictional proposition referring to a fictional object in part estranged from its native fictional context, such as when a fictum is ascribed a relational property relating it to a fictum from a different fictional context, or when a fictum is related to a being from the actual world. Here is an example of an interfictional proposition of the first variety:

$\delta$  First Officer William Riker has a better beard than Luke Skywalker (from episodes VII–IX).

And here’s an example of an interfictional proposition of the second variety:

$\epsilon$  First Officer William Riker has a worse beard than David Lewis.

Indicative interfictional propositions present no serious difficulty and can be accommodated easily enough within Lewisian semantics, given all that has already been said about intrafictionals.

We already know that we can refer to objects from within the context of their native fictions (understood here as worlds of a certain sort) without difficulty. After all, Lewis understands fictional narrative as involving objects from within the actual world successfully referring to objects from within a different world. So, I can refer to  $\text{Riker}_f$  and  $\text{Skywalker}_f$ , even though one is an object in one world and the other is an object in another. Ditto for comparing objects from one world to objects, like Lewis, in the actual world. Now, Lewis understood possible worlds as being spatiotemporally (and causally) isolated domains. Hence, any first-order claim about physical interaction between an object from one world and an object from another is either false or nonsensical. But claims about, e.g., qualitative similarity between objects from different worlds is fine. After all, Lewisian counterpart semantics depend on our ability to coherently speak in this way, since the counterpart relation is determined on the basis of interworld objectual similarity (cf. [7,26,27]). So, to evaluate claims like  $(\delta)$  and  $(\epsilon)$ , we simply abstract the relevant elements from their native domains, combine them into a new constructed (partial) domain, and then evaluate the claim just as we would if we were making a comparative claim between objects from the same world.

#### 5. Ambifictional Propositions

An *ambifictional* is a fictional proposition that is propositionally complex (involving conjunction, disjunction, conditionalization, etc.), where one of the atoms in the semantic molecule refers to a fictional object from within its fictional context, and the other refers to an object from the actual world. The most semantically interesting members of this class are conditionals. Here is an example of an ambifictional proposition:

$\zeta$  If Leonard Nemoy is six feet tall, then Mr. Spock is six feet tall too.

The antecedent of  $(\zeta)$  is an ordinary extensional claim, but the consequent is an intrafictional claim. Hence, the logical form of a proposition like this is ‘ $\varphi \supset f_C(\psi)$ ’: ‘Nemoy is six feet tall  $\supset f_{ST}(\text{Spock is six feet tall})$ ’. N.B. this makes ambifictionals importantly distinct from complex intrafictionals. For example, the statement—if Kirk is six feet tall,

then Spock is six feet tall—is better represented as an intrafictional, by having the fictional operator range over the entire conditional.

Under what conditions is an ambifictional like ( $\zeta$ ) true? Again, no special treatment is required here, given what we have reviewed earlier about intrafictionals. We need only flag that, like ordinary complex indicative propositions, the truth value of ambifictionals would seem to be combinatoric, in the sense that the truth value of the whole is entirely determined and generated using the truth value of its propositional atoms and the functions governing the operative connectives. So, with respect to ( $\zeta$ ), if Nemoy is six feet tall in the actual world, then at the closest possible world *to the actual world* at which the chronicles of Star Trek are true, Spock is six feet tall as well. In other words, the consequent of ( $\zeta$ ) does not change its meaning or determining of truth conditions by being embedded in the material conditional's consequent.

## 6. Intrafictional Counterfactual Propositions

An *intrafictional counterfactual* is a subjunctive fictional proposition with an assumed false antecedent, which ascribes to the fictum some property it would have or might have, given both the consequent's antecedent and the fictum's native fictional context. Here is an example of an intrafictional counterfactual proposition:

$\eta$  Had Spock not sacrificed himself to radiation poisoning (at the end of Star Trek II), the Enterprise might have been destroyed.

Lewis does not discuss this category of proposition. Nonetheless, given ( $\Box \rightarrow / \Diamond \rightarrow$ ) and ( $f_C \star$ ), we once again already have all we need to decipher their complicated semantic. The logical form of a statement like ( $\eta$ ) is ' $f_C(\varphi \Diamond \rightarrow \psi)$ '.

Under what conditions is a proposition like ( $\eta$ ) true? I propose the following order of operations. Supposing ( $\eta$ ) is uttered in the actual world, it is true just in case, at all of the closest worlds to  $w_@$  at which the events comprising the Star Trek narrative are true, at least one of the closest worlds *to them* such that Spock does not sacrifice himself is such that the Enterprise is destroyed. In other words, to evaluate an intrafictional counterfactual, we begin by treating the proposition as though it were an ordinary intrafictional. This will take us to a world at which ordinary intrafictionals about the ficta under examination are true. From there, we then treat the rest of the proposition as though it were an ordinary counterfactual (a might-have counterfactual, in this case), with the exception that, instead of taking  $w_@$  as our base world, we take that (or those) worlds delivered to us via the  $f_C$  operator as our modal base. We thereafter proceed in accordance with ( $\Box \rightarrow / \Diamond \rightarrow$ ) to reach the final evaluation.

## 7. Extrafictional Counterfactual Propositions

An *extrafictional counterfactual* is a subjunctive proposition with an assumed false antecedent that ascribes to a fictional object, qua fictitious artifact, some property that it would or might possess were some antecedent condition to actually occur. Here is an example of an extrafictional counterfactual proposition:

$\theta$  Had the casting director from Star Trek not had a cat named 'Spock,' Spock, the character, might not have been called 'Spock' either.

Under what conditions is a proposition like ( $\theta$ ) true? (N.B., Spock was not actually named after the director's cat). I believe this question has an easy enough answer, though it may seem initially puzzling. To answer it, I will cite in my explanation what I already proposed and defended in §3. I propose that a claim like ( $\theta$ ) is true just in case, among at least one of the nearest possible worlds to  $w_@$  at which the Star Trek casting director does not have a cat named 'Spock', the abstract object,  $\text{Spock}_@$ , is not named 'Spock' either. That is, I take extrafictional counterfactuals to be instances of counterfactual dependency relations between an actual concretum (and its activities) and an actual abstractum (and its abstract, artificial properties). So, this kind of proposition requires no special treatment

outside of  $(\Box \rightarrow)$  and  $(\Diamond \rightarrow)$ . Moreover, since it is an extrafictional, it does not even require that we bring in  $(f_C)$  or  $(f_C \star)$  to evaluate it.

One admittedly odd consequence of this treatment of extrafictionals, and extrafictional counterfactuals more specifically, is that it would entail that it is possible for abstracta to counterfactually depend on concreta. It is one thing to say that the *existence* of an abstractum counterfactually depends on some concretum, but it is quite another to say that the *nature* of an abstractum, what it is like, counterfactually depends on some concretum. But this is precisely what we see with true extrafictional counterfactuals, where what the character is like is largely up to the author. I take it that anyone who has themselves created a fictional character will recognize this truth. During the process of creation, an author usually adds and removes various attributes from the character before settling on some final version. During these emendations, they are likely to say they are working on developing the selfsame character.

## 8. Interfictional Counterfactual Propositions

An *interfictional counterfactual* is a subjunctive proposition with an assumed false antecedent whose antecedent relates ficta from two different fictional contexts, or otherwise relates a fictum with an object from the actual world. Here's an example of an interfictional counterfactual of the first variety:

ι Were Worf, son of Mogh, and Chewbacca to arm-wrestle, Chewie would win.

And here's an example of an interfictional counterfactual of the second variety:

κ Were Worf, son of Mogh, and David Lewis to arm-wrestle, Worf would win.

Undoubtedly, it is by imagining the counterfactual consequences of antecedents of this sort that authors generate believable crossover fictions. The logical form of propositions of this type is on its face no different than intrafictional counterfactual propositions:  $f_C(\varphi \Box \rightarrow \psi)$ . However, a complication results in this case for determining how to assign C. Consider (ι). The context cannot be that of Star Trek, since Chewbacca does not exist in the Star Trek universe. For the same reason, it cannot be that of Star Wars, since Worf does not exist in the Star Wars universe. Presuming statements of this form are not simply vacuously true counterpossibles, it therefore follows that they must be made true by some third context, which includes both ficta in its domain. Such a context may not be given by any known, actual fictional narrative. However, if there already exists a narrative involving a crossover between the two fictions, then it becomes a natural choice at which to frame our evaluations.

Whenever no such crossover fiction exists, I propose that the Lewisian evaluates claims of this sort by constructing a novel context, C, such that C is the most logically consistent union of the domains of the two fictions being combined. Call this "most logically consistent" function on the union ' $\mathcal{O}$ '. So, in this case,  $C = \mathcal{O}(\text{ST} \cup \text{SW})$ , where SW denotes the Star Wars fiction. Hence, we should suppose that (ι) is true just in case at the closest, most logically consistent possible world where both Star Trek and Star Wars are told as known fact, Chewbacca wins the arm-wrestling contest; and (κ) is true just in case at the closest  $\mathcal{O}$ -world where the actual facts of the world exist in conjunction with the Star Trek facts, Worf wins the arm-wrestling contest. I say the Lewisian should take the most logically consistent union of the two fictions since not every two fictions can be combined without contradiction; and it would be counterintuitive to suppose that, e.g., were Worf and Chewbacca or Worf and Lewis to arm-wrestle, some contradictions would be true. This is in fact true of ST and SW, since Star Trek is depicted to exist in the near future, whereas SW, in the distant past, and no moment of any world is simultaneously in the future and the past relative to the present.

## 9. Ambifictional Counterfactual Propositions

We now reach the problematic class: An *ambifictional counterfactual* is a subjunctive conditional with an assumed false antecedent, which relates an actual event in the an-

tedent with a fictional event in the consequent. Here is an example of an ambifictional counterfactual proposition:

- $\lambda$  Had Leonard Nemoy not been six feet tall, Spock, the vulcan, would not have been six feet tall either.

This is the dialectically important class of statement for this essay, so consider two additional ambifictional counterfactuals:

- $\mu$  Had Roddenberry not had a fondness for Sherlock Holmes novels, Commander Data, the android, would not have created a Sherlock Holmes simulation in the holodeck.  
 $\nu$  Had Roddenberry died in 1960, Captain Kirk, commander of the Enterprise, would not have existed.

Statements like ( $\lambda$ ) and the others would appear to have the form  $\phi \Box \rightarrow f_C(\psi)$ . I take it that all of the statements listed above are true. But how to reach that evaluation presents a problem for the Lewisian semantical system.

Consider that the consequent of, e.g., ( $\lambda$ ) is an intrafictional. Hence, ‘Spock’ in its consequent should be understood as referring to  $\text{Spock}_f$ , not  $\text{Spock}_@$ , as in an extrafictional counterfactual of a similar form.  $\text{Spock}_f$  does not exist in  $w_@$ . But Nemoy obviously does exist in  $w_@$ . Hence, if ( $\lambda$ ) is true, it would imply that there is a counterfactual dependence relation obtaining between phenomena of two different worlds—namely,  $\text{Spock}_f$  having a certain height counterfactually depending on Nemoy having that height. But recall that Lewis (1986a) explicitly prohibits such relations in his system. To paraphrase his remarks on this point from §3, to maintain that a proposition like ( $\lambda$ ) obtains is like maintaining that, whereas Spock is six feet tall in  $w_1$  from the perspective of  $w_2$ , he is not from the perspective of  $w_3$ , for which no sense can be made. Hence, it would appear that we must either suppose that all ambifictional counterfactuals are in fact false, appearances to the contrary, or we must reject Lewis’s semantics for either counterfactual discourse or fictional discourse. Lewis’s account is either internally or externally inconsistent.

The paradox, stated more explicitly, runs as follows:

1. Proposition ( $\lambda$ ) is true. (Premise)
2. If ( $\lambda$ ) is true, then the events represented in its consequent counterfactually depend on the events represented in its antecedent. (Premise)
3. The events represented in the antecedent of ( $\lambda$ ) pertain to one world, whereas the events represented in its consequent pertain to another world. (Premise)
4. So, if ( $\lambda$ ) is true, then some events in one world counterfactually depend on events in another. (From (2) and (3).)
5. So, some events in one world counterfactually depend on events in another. (From (1) and (4).)
6. But no events in one world counterfactually depend on events in another. (Premise)
7.  $\perp$ . (From (5) and (6).)

Trekkie lore and pretheoretical intuition justify line (1). Lewis’s semantics for counterfactual discourse justify line (2). Lewis’s semantics for fictional discourse justify line (3). Lewis’s argument against interworld counterfactual dependency justifies line (6). And the inferences to lines (4), (5), and (7) are all deductively valid.

## 10. Revising Lewisian Semantics

For the remainder of the paper, I will consider how a Lewisian would do best to revise Lewis’s system so as to resolve this issue. There would seem to be three salient options for response: the Lewisian could simply reject the problem by rejecting line (1); they could conclude that we should simply scrap Lewis’s system for either counterfactuals or fictionals and create a new one, thus rejecting lines (2) or (3); or they could maintain both systems and instead reject Lewis’s prohibition against interworld counterfactual dependency, thus rejecting line (6). The first two tacks seem to me ill advised, but the third is defensible. I take it that premise (1) is in need of no defense: facts about authors and actors clearly

determine  $C$ , which then determines which intrafictionals are true. I will, however, briefly defend the plausibility of (2) and (3), by way of briefly defending both  $(\Box \rightarrow / \Diamond \rightarrow)$  and  $(f_C)$ .

I say it is ill advised to deny (2). Despite some notable caveats, Lewis's account would still seem to be the best counterfactual semantic available. This is largely on account of its scope. Other theories, like causal interventionist accounts, have difficulty making sense of true/false *acausal counterfactuals*, i.e., nonvacuously true counterfactuals whose antecedent is noncausal or otherwise causally impossible (cf. Bassford & Dolson *forthcoming*). But the Lewisian has no such difficulties. Hence, it seems best to preserve some or another version of Lewisian counterfactual semantics, rather than scrapping them entirely. I say it is also ill advised for a Lewisian to deny (3). (But of course Lewis's theory remains controversial, and there are other accounts on the market—cf., e.g., [28–30].) This is because Lewis supposed that logical space is maximal and generated with an unrestricted principle of propositional combination, such that, if two propositions are logically consistent, then there is a world at which both are true. Hence, the Lewisian already has worlds in the modal domain populated with objects exactly like our fictions. If we can put these worlds and their inhabitants to work within an already intuitively sounding theory of fictional discourse, I am inclined to do so, especially when it comes with no additional ontological cost. Moreover, Lewis's account provides a simple explanation of why it is that both  $(\beta)$  and  $(\gamma)$  are true, even though “Data” does not have two creators. To see why denying (6), and thus rejecting Lewis's ban on interworld counterfactual dependency, is defensible, let us consider in greater detail Lewis's argument against its possibility. I contend that he rejected the possibility of a plausible interworld counterfactual dependence semantic prematurely.

Lewis's ([7]: §1.6, “Isolation”) argument against interworld counterfactual dependence takes the form of a reductio. Suppose, he says, that events,  $E_1$  and  $E_2$ , in two different worlds,  $w_1$  and  $w_2$ , respectively, are counterfactually related to one another. Lewis asks: this counterfactual is supposed to hold *where* exactly? We have four possibilities for making sense of the semantics of claims like  $(\lambda)$ :

- i. If  $E_1$  had not occurred in  $w_1$ ,  $E_2$  would not have occurred in  $w_1$ .
- ii. If  $E_1$  had not occurred in  $w_1$ ,  $E_2$  would not have occurred in  $w_2$ .
- iii. If  $E_1$  had not occurred in  $w_1$ ,  $E_2$  would not have occurred in  $\{w_1, w_2\}$ .
- iv. If  $E_1$  had not occurred in  $w_1$  (in  $\omega_1$ ),  $E_2$  would not have occurred in  $w_2$  (in  $\omega_2$ ).

Lewis rejects (i) because it is not at the world of the antecedent where the occurrence of the consequent is supposed to be relevant, but rather within the world where the consequent is supposed to obtain (cf. [7] pp. 78–79). Moreover, since Spock<sub>f</sub> does not exist in  $w_\oplus$ , this treatment would entail that many true ambifictionals are false. Lewis rejects (ii) because to take this view would commit us to a truth-relativistic modal logic, of the sort he critiqued in passages rehearsed earlier in this essay. But most of us suppose that a truth-absolutist semantic for modal discourse is a better approach. Some fact either obtains or fails to obtain at a world, period. But on this proposal, it may be that, from the perspective of one world,  $w_1$ , some event occurs at  $w_2$ ; but from the perspective of a different world,  $w_3$ , the event does not occur at  $w_2$ . So, we would be committed to thinking that truth in a world is always relative to the perspective of some other world, whose truths are in turn relative to the first. Lewis rejects (iii) too, according to which the counterfactual's consequent should be evaluated at a pair of worlds comprised of  $w_1$  and  $w_2$ . About this proposal, Lewis [7] (pp.79) writes the following:

*“This makes sense, but not I think in a way that could make it true. For I suppose that the closeness of one world-pair to another consists in the closeness of the first worlds of the pairs together with the closeness of the second worlds of the pairs. We have to depart from  $[w_1]$  for the first world of the closest pair, since we have to get rid of  $[E_1]$ . But we are not likewise forced to depart from  $[w_2]$ , for the second world of a closest pair, and what is so close to a world as that world itself? So the second world of any closest pair will be just  $[w_2]$ , at which  $[E_2]$  does occur, so [(iii)] is false.”*

And finally, Lewis rejects (iv) too. He calls it the “one grand world” theory, which makes reference to multiple total domains of possible worlds, categorically symbolized as ‘ $\omega$ ’—i.e., “we think of all the possible worlds as if it were one grand world [ $\omega@$ ], and that starts us thinking that there are other ways it might have been,”  $\omega_1, \omega_2, \omega_3$ , etc. About the grand world proposal for making sense of interworld counterfactuals, Lewis [7] (pp.80) writes the following:

*“[(iv)] is thoroughly misguided. If I am right, the many worlds already provide for contingency, and there is no sense in providing for it all over again. Or else I am wrong, and the many worlds do not provide for genuine contingency. . . . But then it makes no sense to repeat the very method you think has failed, only on a grander scale. The worlds are all the maximal things that are suitably unified. If they fall into grand clusters, and yet grander clusters of clusters, and so on, that is neither here nor there. By ‘worlds’ I still mean all the worlds. (And how could they fall into clusters—what sort of relation could unify a cluster without also merging the worlds within it?) There is but one totality of worlds; it is not a world; and it could not have been different. Therefore, [(iv)] is nonsense, intelligible only if taken as vacuous.”*

Hence, Lewis concludes, we should simply deny all interworld counterfactuals. Lewis’s critiques of (i), (ii), and (iii) strike me as plausible, but I believe that his critique of proposal (iv) is far too hasty, and that it is in fact precisely in this way the Lewisian should revise the semantic theory to accommodate true ambifictional counterfactuals. I will offer Lewis’s critique of (iv) three replies.

Reply 1. Lewis’s critical argument against (iv) takes a disjunctive eliminative form. He reasons the following: either his modal semantic provides for contingency or it does not. If it does, then there is no need to further complicate the picture of logical space. If it does not, then adding complication by creating a second-order set of worlds will not account for it either, since it is the same process by which the first-order domain was accounted. In that case, we ought to accept the initial methodology as applied to the first-order domain as already sufficient. So, either way, there is no need to complicate his picture of logical space. I take it that Lewis’s argument is valid but unsound. The second horn in the dilemma is blunt. His account certainly provides for *some* contingency. We can accept this while nevertheless denying that it provides for *all* of it. Ambifictional counterfactuals, in fact, may push us into seeing that the model, as it currently stands, is insufficient and does not yet have the breadth to represent all of the contingency structuring the cosmos.

Reply 2. Again, with respect to Lewis’s critical argument, he reasons that, if a plurality of worlds does not already account for all contingency, then a plurality of grand worlds will not do it either. But the methodological principle undergirding Lewis’s argument here is no good. Consider the history of extensional logic. We developed ways of reasoning about the first-order domain, in which we treat objects in the world as subjects, and we treat properties as predicates. We found, however, that this model on its own was insufficient to capture all extensional truths, such as truths about concepts and properties themselves. ‘Red is my favorite color’, for example, is a very different statement than ‘All red things are my favorite color’. So, to accommodate these sorts of proposition into our semantic theory, we developed second-order extensional logic, whereby we now treat properties as primitive subjects in the calculus, and we treat properties of properties as their predicates. Many are satisfied that this iterative process was well founded and secured for us greater semantic understanding. Just so, we would do best to use what has already worked in extensional logic and just apply it, in this case, to intensional logic too, thereby developing a second-order modal logic. I believe this would allow us to capture the remaining modal facts, which are apparently not currently captured in their entirety in Lewis’s first-order modal logic.

Reply 3. Lewis concludes his argument with a challenge to the grand world semantic: state some principle by which worlds might cluster into unified galaxies, complete with their own isolated modal facts. He does not believe there could be any plausible principle. Suppose there are multiple modal domains. So, e.g., relative to one modal galaxy,  $\omega_1$ ,

$R(w_1, w_2)$ , but relative to another,  $\omega_2$ ,  $\sim R(w_1, w_2)$ . Call the first  $w_1$  ' $w_{1\omega_1}$ ', and the second ' $w_{1\omega_2}$ '. This causes no problem; after all, we already do this when we consider modal facts about objects and properties—we imagine that something answering to their names within a different domain, complete with its own facts, which may or may not correspond with our own, is a certain way. The suggestion is only that we do the same thing with worlds and ask what the complete modal galaxy,  $\omega_@$ , may or may not have been like. We might even develop a corresponding intergalactic modal logic and pose logically and metaphysically rich questions about the new model, such as whether worlds from different grand worlds are identical or only bear counterpart relations to one another. One such interesting question is the one Lewis asks: what unifies grand worlds into isolated clusters? One answer, given what has already been said here, is that modal galaxies could be unified by interworld counterfactual dependencies. All such dependencies are non-causal, so Lewis's principle of the causal isolation of possible worlds remains unviolated. Again, ambifictional counterfactuals push us to think that there could be such interworld counterfactual dependencies. Perhaps there are also instances of interworld counterfactual dependencies besides those involving ficta. If so, then those instances might configure into the principle logically isolating and unifying modal galaxies as well.

For these reasons, I believe that, of the four responses to the paradox of ambifictional counterfactuals for Lewis's semantics, the Lewisian would be best advised to expand their semantic by integrating modal galaxies (grand worlds) into their theory's domain, and then use them to make sense of how interworld counterfactual dependence is possible. This amounts to the suggestion that we evaluate counterfactuals like ( $\lambda$ ) by supposing that at the nearest modal galaxy to  $\omega_@$  where Nemoy is not six feet tall at  $w_@$ , Spock<sub>f</sub> is not six feet tall (in his worlds) either. Lewis's critique of the grand world model of logical space is unsound, and the revised Lewisian model would introduce a host of new, interesting philosophical questions, semantical and otherwise.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Roddenberry, G. (creator); *Star Trek: The Original Series*; NBC: New York, NY, USA, 1966.
2. Roddenberry, G. (creator); *Star Trek: The Next Generation*; Paramount Pictures: Los Angeles, CA, USA, 1987.
3. Lewis, D.K. *Counterfactuals*; Wiley-Blackwell: Hoboken, NJ, USA, 1973.
4. Lewis, D.K. Counterfactuals and Comparative Possibility. *J. Philos. Log.* **1973**, *2*, 418–446. [CrossRef]
5. Lewis, D.K. Counterfactual Dependence and Time's Arrow. *Nous* **1979**, *13*, 455–476. [CrossRef]
6. Lewis, D.K. Ordering Semantics and Premise Semantics for Counterfactuals. *J. Philos. Log.* **1981**, *10*, 217–234. [CrossRef]
7. Lewis, D.K. *On the Plurality of Worlds*; Wiley-Blackwell: Hoboken, NJ, USA, 1986.
8. Lewis, D.K. Postscripts to 'Counterfactual Dependence and Time's Arrow'. In *Philosophical Papers*; Oxford University Press: Oxford, UK, 1986; pp. 52–66.
9. Nute, D. Counterfactuals and Similarity of Worlds. *J. Philos.* **1975**, *72*, 773–778. [CrossRef]
10. Stalnaker, R.C. A Theory of Conditionals. In *Studies in Logical Theory 2*; Rescher, N., Ed.; Wiley-Blackwell: Hoboken, NJ, USA, 1968; pp. 98–112.
11. Arregui, A. On Similarity in Counterfactuals. *Linguist. Philos.* **2009**, *32*, 245–278. [CrossRef]
12. Bassford, A.D.; Dolson, C.D. Counterfactual Similarity, Nomic Indiscernibility, and the Paradox of Quidditism. *Inquiry* **2019**. [CrossRef]
13. Kment, B. Counterfactuals and Explanation. *Mind* **2006**, *115*, 261–310. [CrossRef]
14. Starr, W. Counterfactuals. In *The Stanford Encyclopedia of Philosophy*; Zalta, E.N., Ed.; Stanford University: Stanford, CA, USA, 2019; Available online: [plato.stanford.edu](https://plato.stanford.edu) (accessed on 16 September 2023).
15. Lewis, D.K. Truth in Fiction. *Am. Philos. Q.* **1978**, *15*, 37–48.
16. Blumberg, K.; Holguin, B. *Settling Fictions*; University of Texas at Austin: Austin, TX, USA, 2023.
17. Friend, S. Fictional Characters. *Philos. Compass* **2007**, *2*, 141–156. [CrossRef]

18. Kroon, F.; Voltolini, A. Fiction. In *The Stanford Encyclopedia of Philosophy*; Zalta, E.N., Ed.; Stanford University: Stanford, CA, USA, 2019; Available online: [plato.stanford.edu](https://plato.stanford.edu) (accessed on 16 September 2023).
19. Kroon, F.; Voltolini, A. Fictional Entities. In *The Stanford Encyclopedia of Philosophy*; Zalta, E.N., Ed.; Stanford University: Stanford, CA, USA, 2023; Available online: [plato.stanford.edu](https://plato.stanford.edu) (accessed on 16 September 2023).
20. Rast, E. Classical Possibilism and Fictional Objects. In *Truth in Fiction*; Lihoreau, F., Ed.; Transaction Books: Piscataway, NJ, USA, 2011; pp. 77–92.
21. Walton, K. Fearing Fictions. *J. Philos.* **1978**, *75*, 5–27. [[CrossRef](#)]
22. Eagleton, T. *Literary Theory: An Introduction*; University of Minnesota Press: Minneapolis, MN, USA, 2008.
23. Berto, F. Modal Meinongianism for Fictional Objects. *Metaphysica* **2008**, *9*, 205–218. [[CrossRef](#)]
24. Falguera, J.L.; Martínez-Vidal, C.; Rosen, G. Abstract Objects. In *The Stanford Encyclopedia of Philosophy*; Zalta, E.N., Ed.; Stanford University: Stanford, CA, USA, 2022; Available online: [plato.stanford.edu](https://plato.stanford.edu) (accessed on 16 September 2023).
25. Lewis, D.K. New Work for a Theory of Universals. *Australas. J. Philos.* **1983**, *61*, 343–377. [[CrossRef](#)]
26. Lewis, D.K. Counterpart Theory and Quantified Modal Logic. *J. Philos.* **1968**, *65*, 113–126. [[CrossRef](#)]
27. Lewis, D.K. Counterparts of Persons and Their Bodies. *J. Philos.* **1971**, *68*, 203–211. [[CrossRef](#)]
28. Berto, F. Modal Meinongianism and Fiction: The Best of Three Worlds. *Philos. Stud.* **2011**, *152*, 313–334. [[CrossRef](#)]
29. Sainsbury, R.M. Fiction and Acceptance-Relative Truth, Belief and Assertion. In *Truth in Fiction*; Lihoreau, F., Ed.; Transaction Books: Piscataway, NJ, USA, 2011; pp. 137–152.
30. Walton, K. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*; Harvard UP: Cambridge, MA, USA, 1990.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.