

Counteressential Conditionals

Kenneth L. Pearce
Valparaiso University

Abstract

Making sense of our reasoning in disputes about necessary truths requires admitting non-vacuous counterpossibles. One class of these is the counteressentials, which ask us to make contrary to fact (and therefore contrary to possibility) suppositions about essences. A popular strategy in accounting for non-vacuous counterpossibles is to extend the standard possible worlds semantics for subjunctive conditionals by the addition of impossible worlds. A conditional $A \Box \rightarrow C$ is then taken to be true if all of the nearest A worlds (whether possible or impossible) are C worlds. I argue that a straightforward extension of the standard possible worlds semantics to impossible worlds does not result in a viable account of counteressentials and propose an alternative covering law semantics for counteressentials.

Keywords: conditionals; essences; possible worlds; grounding; hyperintensionality

1 The Euthyphro Dispute

Suppose that God exists necessarily and essentially possesses a character which ensures that God eternally disapproves of torture. Suppose further that it is a necessary truth that God eternally disapproves of all and only those actions that are morally wrong. Within the framework of these suppositions, it is possible for the Euthyphro dispute to arise. One disputant (call her ‘Thea’) may hold that torture is wrong because God disapproves of it, while another disputant (call him ‘Deon’) may hold that God disapproves of torture because it is wrong.¹ The ‘because’ in these claims is hyperintensional: the disputants agree that the two sentences joined by ‘because’ express necessary truths, but disagree on the relation between those necessary truths.

Thea’s ‘because’ claim commits her to the following conditional:

1. I use ‘disapproves of’ rather than ‘prohibits’ because Thea should not be understood as holding that God at some time issues a prohibition against torture and thereby brings it about that (beginning from that time) torture is wrong. Instead, it is God’s eternal attitude, following necessarily from God’s character, which Thea takes to ground the wrongness of torture.

Thea's Conditional: If God did not disapprove of torture, torture would not be wrong.

Deon's 'because' claim commits him to a different conditional which is apparently inconsistent with Thea's:

Deon's Conditional: If God did not disapprove of torture, torture would still be wrong.

Thea and Deon agree that the antecedent which is shared by these two conditionals is necessarily false due to the divine essence. Intuitively, however, it seems that Thea's 'because' claim entails that her conditional is true and Deon's is false, while Deon's 'because' claim entails that his conditional is true and Thea's is false. Furthermore, intuitively, the two conditionals are inconsistent with one another.

The Euthyphro dispute is merely one vivid example of a much more widespread phenomenon: reasoning about what is necessarily true frequently requires us to make impossible suppositions and these suppositions and the results we take them to have are frequently expressed as counterpossible conditionals (see Nolan 1997, §2; 2013, §2.2). The conditionals involved in the Euthyphro dispute are what I will call *counteressentials*: conditionals that ask us to consider contrary to fact, and therefore contrary to possibility, suppositions about essences. In the example, we are asked to make suppositions which the disputants take to be contrary to the divine essence. However, these sorts of conditionals are not found only in philosophy of religion. For instance, it is very plausible to suppose that chemical substances have their molecular composition essentially, but in order to determine what this molecular composition is it is important to consider how the experiments would come out on various assumptions about that composition. All but one of these suppositions will be necessarily false.

Any adequate theory of subjunctive conditionals must make sense of the use of counterpossibles – and specifically counteressentials – in reasoning and disputing about necessary truths. The most straightforward and intuitive way of doing this would be to allow that counterpossibles may differ from one another in their truth values and entailment relations so that, for instance, Thea's claim that torture is wrong because God disapproves of it entails the truth of her conditional and the falsity of Deon's. However, the standard possible worlds semantics for subjunctive conditionals cannot permit this. Instead, it makes all counterpossibles true vacuously.

There have been two main lines of response to this difficulty. The first, which was influentially advocated by David Lewis, accepts the result that all counterpossibles are vacuous and attempts a pragmatic, rather than semantic, explanation of why some counterpossibles can be felicitously asserted in a given context and others cannot (Lewis 1973, §1.6). On this view, Thea and Deon both ought to recognize that both conditionals are true; however, their background theories somehow give rise to a difference in assertibility.

The second approach is to introduce impossible worlds in order to generate a semantic distinction that will make some counterpossibles (non-vacuously) true and others false.

In this paper, I will assume (contrary to Lewis) that the Euthyphro dispute and other disputes like it do indeed involve or imply a disagreement about the truth values of counteressential conditionals. Having made this assumption, my aim will be to understand how these conditionals are to be evaluated and, in particular, how they are related to the hyperintensional ‘because’ claims that give rise to them. I will begin by arguing that the impossible worlds approach fails. I will then develop an alternative covering law semantics for counteressentials. My approach will give central importance to the relationship between these conditionals and the hyperintensional ‘because’. Finally, I will consider whether a modified impossible worlds analysis might be able to achieve the same results as mine. Such an analysis, I will show, faces very serious difficulties.

2 Against Impossible Worlds

The most obvious way of adapting the possible worlds semantics to accommodate non-vacuous counterpossibles is to introduce impossible worlds (Nolan 1997, §4; 2013, 363). Such an approach would retain the basic idea that a subjunctive conditional $A \Box \rightarrow C$ is true if and only if all the nearest A-worlds are C-worlds and simply allow that in some cases the nearest A-worlds may be impossible worlds.

The most pressing problem for such a semantics is to give an account of nearness. Nearness cannot simply be overall similarity, since, as Kit Fine pointed out, it is true that *if Nixon had pressed the button there would have been a nuclear holocaust*. Yet, on the assumption that there will not be a nuclear holocaust, the world at which the button malfunctions is more similar to the actual world than the one at which the button works, thus apparently rendering the conditional false (Fine 1975, 452).

Once impossible worlds are introduced, this objection becomes even more pressing. For any false proposition you like, there will be an impossible world at which that proposition is true and everything else is the same – including the truth of that proposition’s negation. Thus not only will we apparently get the result that *if Nixon had pressed the button, no nuclear holocaust would have occurred*; we will also get the result that *if Nixon had pressed the button, Nixon wouldn’t have pressed the button*.

David Lewis responded to Fine’s objection by developing a priority ordering of respects of similarity designed to explain why some worlds that have seemingly momentous changes (e.g., future nuclear holocausts) should, in most contexts, be regarded as nearer than some worlds having seemingly minor changes (e.g., disappearing electrons in the launch mechanism). Crucially, according to Lewis, “It is of the first importance to avoid big, widespread, diverse violations of [natural] law” (Lewis 1979, 472). Note, however, that Lewis here relies on the notion of *magnitude* of violation of the laws of nature. In some cases, the magnitude of a violation can actually be quantified as if, for instance, we consider *how much* faster than the speed of light a particle is supposed to move.

In many other cases, though, we have to rely on our intuitions about how ‘big’ the violation is. As we get further from actuality, these intuitions get murkier and our judgments get quite fuzzy.

Attempts to adapt this approach to impossible worlds would, again, make things even worse, for many of the impossible scenarios we need to consider are simply beyond our powers of conception. Of course there are some cases in which we can easily say that one violation of necessary truth is bigger than another – as, for instance, the supposition $2+2 = 8$ involves a bigger violation of necessary truth than the supposition $2+2 = 5$. But how should Thea compare the impossible world at which God does not disapprove of torture and torture is still wrong with the world at which God does not disapprove of torture and torture is not wrong? Here there is a collision between normative ethical claims and meta-ethical claims, both of which are taken to be necessary truths. As long as we are trafficking in impossibilities, perhaps we might even consider the impossible world at which God both disapproves and does not disapprove of torture and torture is both wrong and not wrong.

What this shows is that Lewis’s own account of nearness in terms of similarity cannot be straightforwardly adapted by the proponent of impossible worlds. These proponents are of course free to propose alternative accounts of nearness for the evaluation of counterpossibles. However, there is a simpler approach available. It is our understanding of the hyperintensional relation (expressed by the word ‘because’) which Thea alleges to exist between divine disapproval and wrongness that makes us judge that *if Thea’s theory were correct and God did not disapprove of torture, then torture would not be wrong*. Taking this connection between ‘because’ and ‘if’ as our starting point will enable us to develop a simple and satisfactory account of counteressential conditionals.

3 A Covering Law Semantics For Counteressentials

The leading alternative to the worlds approach to evaluating subjunctive conditionals takes some relevant state of affairs smaller than an entire world and modifies it to render the antecedent of the conditional true. The conditional is then taken to be true if the consequent becomes true as a ‘result’ or ‘outcome’ of the modification. These are sometimes called ‘covering law’ approaches because the consequent is usually said to be a result or outcome of the antecedent if and only if it follows according to some law. Approaches along this general line have been advocated by Tim Maudlin and Kit Fine (Maudlin 2007, ch. 1; Fine 2012b).

If this approach is to succeed, we must develop a procedure which will allow us, for each conditional, to identify the relevant state of affairs and determine what modification should be made. Further, we must define the relevant notion of ‘result’ or ‘outcome.’ I will show how to do this for counteressentials.

A counteressential asks us to make a contrary to fact (and therefore contrary

to possibility) supposition about some particular essence. This essence is the state of affairs we should consider in evaluating the conditional.² We then consider every way of modifying that essence to make the antecedent true, and ask whether the consequent is an outcome of every such modification. Accordingly, Thea's Conditional is to be understood as claiming that, in whatever way we modify (in thought) the divine essence to render it consistent with God's failing to disapprove of torture, we will get the result that torture is not wrong.

It is important that at this stage we do *not* consider modifications of the essences of torture or wrongness. That it is God, and not (e.g.) wrongness, that is to be modified is indicated by God's appearing in subject position. Additional essences may perhaps be modified in the course of determining the result, but this modification should not be part of our initial supposition.

Having identified the state of affairs and the modifications to that state of affairs to be considered, we must next determine what the result relation should be. It certainly cannot be a causal relation, for essences are not generally (efficient) causes. Instead, I suggest, the relation in question is a *grounding* relation.

Consider, by way of analogy, Maudlin's account of subjunctive conditionals with physically possible antecedents. According to Maudlin, the laws of nature are Laws of Temporal Evolution (LoTEs) that tell us how to evolve a system forward or backward in time. For most counterfactuals, the relevant state of affairs is the state of the universe at a time.³ One then uses the relevant LoTEs to evolve the system in the appropriate direction and see if the consequent comes out true.

Similarly, many metaphysicians hold that there are metaphysical laws of grounding which govern the manner in which less fundamental things arise from more fundamental things (see, e.g., Fine 2011, 75-76). These laws generate a hierarchy of fundamentality (44-45, 51). My suggestion is that we use these metaphysical laws for evaluating counteressentials analogously to Maudlin's use of physical laws for evaluating counterfactuals. Thus we first modify the essence, then use the general facts about grounding to determine whether the modification of the essence requires the consequent to come out true. Since the grounding relation is hyperintensional, this will get the desired result that necessary equivalents *cannot* be substituted into the antecedent or consequent of a counteressential conditional, and hence that conditionals with necessarily false antecedents and consequents may differ in their truth-values.⁴

On this view, the Euthyphro dispute comes out just the way it should, as a dispute about the grounds of wrongness. The hyperintensional 'because' which is at issue here expresses a grounding relation. Thea holds that wrongness is grounded in divine disapproval. Accordingly, on her view, the supposition that

2. If essences are not states of affairs, then in evaluating the Euthyphro conditionals we should consider the state of affairs *God's essence being such and such*.

3. Various complications, which Maudlin addresses in detail, are introduced by the theory of relativity.

4. Fine 2012a argues at length that substitution of logical equivalents is invalid even for counterfactuals with possible antecedents.

God's essence is such that God does not disapprove of torture will have the consequence that torture is not wrong. Deon holds that wrongness is metaphysically prior to divine disapproval so that, on the supposition that God did not disapprove of torture, God would be in disagreement with the moral facts. The covering law approach succeeds where the worlds approach fails in providing a satisfactory account of counteressentials.

4 Impossible Worlds Restored?

My strategy has been to adapt the covering law semantics for counterfactuals to the case of counteressentials by introducing metaphysical laws that play a role analogous to that played by natural laws. As we have seen, natural laws also play an important role in Lewis's possible worlds semantics. Could the same metaphysical laws be employed to save the impossible worlds approach to counteressentials?⁵

Lewis employs laws of nature to get the result that, in evaluating counterfactuals, some differences in the contingent features of worlds (those that involve violations of natural laws) matter more than others. The analogous move would be to say that some differences in the non-contingent features of worlds (those that involve violations of metaphysical laws) matter more than others. The proponent of impossible worlds will likely wish to introduce a priority ordering of laws, so that we first minimize violations of logical/mathematical laws, then metaphysical laws, then natural laws.

This approach gets the right result from Thea's perspective: the world at which God both approves and disapproves of torture and torture is both wrong and not wrong violates logical laws, and so is farther from actuality than any world that violates only metaphysical laws. The world at which God does not disapprove of torture and torture is still wrong violates the metaphysical law about the grounds of moral truths. The world at which God does not disapprove of torture and torture is not wrong is still an impossible world, since it contains a violation of the divine essence, but it doesn't violate any metaphysical laws. Hence, it is the closest of the three. Accordingly, given Thea's views about the grounds of moral truths, Thea's Conditional is true and Deon's Conditional is false.

From Deon's perspective, things are not so simple. On the present proposal, Deon's evaluation of the conditionals will depend on his positive meta-ethical position, which we have so far left unspecified. To see this, suppose Deon is a meta-ethical Platonist who holds that wrongness is an objective, primitive, categorical property which is an essential feature of certain action types. So just as it is essential to torture that it involves causing (physical or psychological) suffering it is also essential to torture that it is wrong. At the impossible world where God disapproves of torture and torture is still wrong, there is a violation of a metaphysical law, the one that grounds God's disapproval in wrongness. But there is another impossible world at which God does not disapprove of torture

5. I thank an anonymous referee for raising this question.

and the essence of torture does not include wrongness. At this world, there are only changes to essences, and no violations of metaphysical law. According to the present proposal, the second world will be nearer to actuality than the first, with the result that Deon should endorse Thea's Conditional and reject his own.

On the present proposal, Deon's evaluation of the conditionals depends on whether, according to his view, the supposition that torture is not wrong involves a violation of metaphysical law or only a change in essences. Intuitively, however, Deon's evaluation of the conditionals does *not* depend on this: as soon as Deon commits himself to the claim that wrongness is metaphysically prior to divine disapproval, he is thereby committed to endorsing his own conditional and rejecting Thea's, regardless of his positive meta-ethical views.

The worlds theorist cannot hold that violations of metaphysical law universally trump differences to essences in evaluating the nearness of worlds. Yet she must hold that such violations *sometimes* trump or else she will not correctly capture the implications of Thea's view. It is perhaps natural for the worlds theorist to hold that this is one of the many cases in which there is vagueness about which worlds are most similar. As Lewis, says, "Different resolutions of the overall vagueness of overall similarity are appropriate in different contexts" (Lewis 1979, 465). However, it is not at all clear that the worlds theorist can use this vagueness to her advantage. It seems that, in the very same conversational context, Thea ought to regard her conditional as determinately true and Deon's as determinately false, while Deon ought to regard his conditional as determinately true and Thea's as determinately false. Despite what Lewis calls "the extreme shiftiness and context-dependence of similarity," (466), it is implausible to suppose that a different similarity metric should be employed for evaluating these two very similar conditionals in the very same context.

Note further that impossible worlds were introduced here precisely in order to secure a difference that was semantic, rather than merely pragmatic. This does not mean that the impossible worlds theorist cannot appeal to contextual or pragmatic factors here; the role of such factors in determining the appropriate nearness metric is a central feature of the entire worlds-based approach. The challenge facing the impossible worlds theorist is rather that she must identify contextual or pragmatic factors which give rise to a difference in the truth values and entailments of the conditionals, and she must do this in a way that makes sense of the Euthyphro dispute as a substantive debate about the relationship between God and morality.⁶ I am not able to provide any general argument for the conclusion that this cannot possibly be done, but the identification of the appropriate metric of nearness in the context of the Euthyphro dispute and other similar disputes presents a formidable challenge for the impossible worlds theorist and it is a challenge which has yet to be met.

6. It is at this point that the analysis proposed by Brogaard and Salerno 2013 falters. Brogaard and Salerno employ a 'partly epistemic' notion of relevant a priori implication rather than talking about essences or metaphysical laws of grounding, as on the proposal considered here. This has the effect of making the Euthyphro dispute a verbal dispute between individuals employing different concepts with different a priori implications rather than a dispute about the (alleged) objective metaphysical relation between God and wrongness.

The covering law approach holds that it is a mistake to look for ‘nearness’ or ‘overall similarity’ of worlds in the first place (see Maudlin 2007, 32-33). Instead, we should follow out the laws from the differences specified in the antecedent to see whether the consequent results. This gets the right evaluations without any *ad hoc* maneuvers.

Nevertheless, let us suppose for the sake of argument that the worlds theorist somehow manages to give a non-*ad hoc* account of the relative weight to be accorded to changes in essence and violations of metaphysical law in the context of the Euthyphro dispute. Even so, the worlds theorist will now be committed to regarding the hyperintensional metaphysical laws as prior to the nearness relation. This being so, it seems unlikely that impossible worlds will provide a reductive account of hyperintensionality as some proponents have hoped they would (see, e.g., Nolan 2013, 366; Brogaard and Salerno 2013).

Lewis believed that he could use his plurality of worlds to give a reductive analysis of apparently intensional phenomena. Crucial to this project was Lewis’s view that each world is a ‘Humean mosaic’ – a distribution of categorical properties through space and time – and thus is fundamentally purely extensional (Lewis 1986, 70-86; 1994). Apparently intensional phenomena are to be analyzed by relations between or comparisons among worlds. So, to use Lewis’s example, the concepts *having kidneys* and *having a heart* are coextensive across the actual world. Their difference in meaning can be accounted for by considering the fact that, when allowed to range over possibilities, the terms *do* differ in extension (Lewis 1986, §1.5). Subjunctive conditionals, since they are not truth functional, are among the intensional phenomena to be reduced. Accordingly, Lewis must (and does) provide a non-circular account of the truth conditions of subjunctive conditionals purely in terms of the extensional/categorical features of worlds. This involves giving a purely extensional (this-worldly) account of the laws of nature employed in our judgments of nearness (Lewis 1973, §3.3).

This, however, points to a serious disadvantage of Lewis’s approach: on Lewis’s own view, the proper evaluation of counterfactuals requires that we have a prior and independent analysis of laws of nature. However, as the covering law approach shows, once we have such an analysis, we are in a position to evaluate counterfactuals without appeal to possible worlds or nearness relations between them. On the view we are now considering, essences and metaphysical laws of grounding would both play important roles in evaluating the similarity of worlds. Accordingly, the envisaged reduction requires a prior and independent analysis of essences and metaphysical laws. But once we have such an analysis we are in a position to evaluate counteressential conditionals without appealing to impossible worlds at all. The same point would of course apply if we gave up on the reductive project and took laws and/or essences as primitive.

The addition of impossible worlds to a basically Lewisian semantics for subjunctive conditionals does not result in a satisfactory account of counteressential conditionals such as those involved in the Euthyphro dispute. Yet making sense of these kinds of disputes about necessary truths is one of the key motivations for positing impossible worlds in the first place. Unless and until well-motivated solutions to the problems I have outlined are identified, the need to make sense

of the role of counterpossibles in reasoning about necessary truths should not be regarded as a reason for positing impossible worlds, but rather as a reason for rejecting the worlds-based approach to subjunctive conditionals. The covering law semantics – supplemented with metaphysical laws of grounding – succeeds where impossible worlds fail.⁷

References

- Brogaard, Berit, and Joe Salerno. 2013. “Remarks on Counterpossibles.” *Synthese* 190 (4): 639–660.
- Fine, Kit. 1975. Critical Notice of *Counterfactuals* by David Lewis. *Mind* 84 (335): 451–458.
- . 2011. “Guide to Ground.” In *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Schnieder, 37–80. Cambridge: Cambridge University Press.
- . 2012a. “A Difficulty for the Possible Worlds Analysis of Counterfactuals.” *Synthese* 189 (1): 29–57.
- . 2012b. “Counterfactuals Without Possible Worlds.” *Journal of Philosophy* 109 (3): 221–246.
- Lewis, David. 1973. *Counterfactuals*. Malden, Mass.: Basil Blackwell, Ltd.
- . 1979. “Counterfactual Dependence and Time’s Arrow.” *Noûs* 13 (4): 455–476.
- . 1986. *On the Plurality of Worlds*. Malden, MA: Blackwell.
- . 1994. “Humean Supervenience Debugged.” *Mind* 103 (412): 473–490.
- Maudlin, Tim. 2007. *The Metaphysics Within Physics*. Oxford: Oxford University Press.
- Nolan, Daniel. 1997. “Impossible Worlds: A Modest Approach.” *Notre Dame Journal of Formal Logic* 38 (4): 535–572.
- . 2013. “Impossible Worlds.” *Philosophy Compass* 8 (4): 360–372.

7. I thank Sandra Visser, an anonymous referee, and the editors of *Thought* for comments on previous drafts.