Counterfactual Logic and the Necessity of Mathematics

Samuel Z. Elgin

March, 2020

Abstract

This paper is concerned with counterfactual logic and its implications for the modal status of mathematical claims. It is most directly a response to an ambitious program by Yli-Vakkuri and Hawthorne (2018), who seek to establish that mathematics is committed to its own necessity. I claim that their argument fails to establish this result for two reasons. First, their assumptions force our hand on a controversial debate within counterfactual logic. In particular, they license counterfactual strengthening—the inference from 'If *A* were true then *C* would be true' to 'If *A* and *B* were true then *C* would be true'—which many reject. Second, the system they develop is provably equivalent to appending Deduction Theorem to a **T** modal logic. It is unsurprising that the combination of Deduction Theorem with **T** results in necessitation; indeed, it is precisely for this reason that many logicians reject Deduction Theorem in modal contexts. If Deduction Theorem is unacceptable for modal logic, it cannot be assumed to derive the necessity of mathematics.

Introduction

Mathematical truths necessarily obtain.[1] While it is possible for Hillary Clinton to have won the 2016 presidential election, it is necessary that $2 + 2 = 4$; while the Axis powers could have won World War II, it could not be that negative numbers have real square roots; and while there are some possible worlds in which there are an even number of stars, there are none in which all Fermat numbers are prime. History might have progressed far differently than it actually did, and the laws of physics might even have diverged wildly from what they actually are, but, the received wisdom goes, pure mathematics concerns what is necessarily true—it may even be the paradigmatic example of a realm of necessary truths.

[1] I would like to thank the attendees of the 13th Annual Cambridge Graduate Conference on the Philosophy of Logic and Mathematics as well as the Metaphysics and Semantics Working Groups at the University of California, San Diego for their feedback on earlier versions of this paper, as well as Alex Roberts for his helpful comments and Juani Yli-Vakkuri for his illuminating email correspondence.

This much is uncontroversial (or, at least, as uncontroversial as anything ever is in philosophy), but there is currently no consensus on the foundations for the necessity of mathematics. In virtue of what do these truths, rather than others, hold necessarily? Are we justified in our collective confidence that they could not have been otherwise? Is there a division of labor, such that mathematics provide the truths and philosophy the necessity, or is mathematics itself committed to the necessity of its claims?

Numerous proposals are available in the literature. According to one, the necessity of mathematics is secured by the strength of our intuitions.[2] Perhaps conceivability is a guide to possibility; the fact that it is conceivable that p is evidence that it is possible that p, and the fact that it is inconceivable that p is evidence that it is impossible that p. If so, then our inability to conceive of a way for two and two to make five is evidence that it is impossible for two and two to make five. And if all mathematical falsehoods are similarly inconceivable, we can be confident in the necessity of mathematical truths. Of course, this strategy does not determine the metaphysical basis for the necessity of mathematics, but it could explain why our belief in that necessity is justified. Alternatively, according to neologicists—who maintain that arithmetic is reducible to logic—the necessity of mathematics results from the necessity of logic.[3] Arguably, the necessity of logic is as reasonable a starting-point as any in modal inquiry, so if arithmetic is reducible to logic, then logical truths generate arithmetic truths that necessarily obtain. However, in light of Gödel's incompleteness theorem, neologicists typically aim only to establish the necessity of a fragment of mathematics.[4] Still others argue that we ought not be nearly so confident in the necessity of mathematics as we currently are.[5] Mathematicians are standardly content to prove that something is true; they seldom bother to prove that it is

[2]See Bealer (2002). For a more general discussion of the connection between conceivability and possibility (especially in light of the Kripke (1980) development of the necessary *a posteriori*) see Gendler and Hawthorne (2002).

[3]See, e.g., Hale and Wright (2001).

[4]Yli-Vakkuri and Hawthorne, for example, claim "The neologicist strategy has inherent limitations. It can, at best, establish only the necessity of those mathematical truths that are provable in whatever axiomatic system it uses. By Gödel's first incompleteness theorem, we know that these cannot even include all truths of first-order arithmetic" (pg. 4). I find this modesty premature. It is worth recalling, as philosophers are often prone to forget, that arithmetic is only incomplete on the assumption that its axioms ought to be decidable—i.e., that an infinitely large computer with an infinite amount of time ought to be able to determine whether a given formula is an axiom. There are numerous complete, albeit undecidable, axiomatizations of arithmetic. Whether decidability is an appropriate restriction depends largely on our theoretical aims. I see no reason why axioms ought to be decidable when the subject is the reduction of arithmetic to logic; all that is required is that each axiom be a principle of logic. For example, the ω-rule, according to which one may infer $\forall x Fx$ after infinitely many steps determining that $Fa, Fb, ...$ is undecidable but arguably a principle of logic (minimally, it seems as plausibly a principle of logic as Hume's Principle, according to which the number of Fs = the number of Gs just in case there is a one-to-one correspondence between the Fs and Gs, something neologicists often assume). I suspect that this humility arises because neologicists are typically committed not only to the reduction of arithmetic to logic in general, but to Frege (1884)'s derivation in particular. This strategy inevitably inherits the incompleteness of Peano arithmetic.

[5]See, e.g., Hodges (Forthcoming).

necessarily true. Indeed, terms like 'necessity' and 'possibility' are conspicuously absent from the vast majority of mathematical texts. Philosophers, some claim, step in when mathematicians' work is complete and (perhaps erroneously) attribute necessity to the results of their theorems.

Recently, Yli-Vakkuri and Hawthorne (2018) provide a novel defense for the necessity of mathematics. They argue that counterfactual logic and mathematical practice jointly entail that mathematics is committed to its own necessity: that, for any sentence *S* within the language of pure mathematics, if *S* is true then *S* is necessarily true. Their assumptions do not merely entail that mathematics is committed to the necessity of its claims, but to an **S5** modal logic in particular. Its modal commitments run deep.

When I first encountered this paper, I was captivated by its result. It seemed to me that—at long last—we had no need to rely on the strength of intuition or the dubious program of neologicism. A rigorous derivation could take their place. Perfectly innocuous assumptions about counterfactual logic entail that mathematics is committed to its own necessity. Indeed, I suspected that this would eventually be seen as one of the most significant contributions to the philosophy of mathematics in many years.

My doubts have since developed. I no longer believe that this program succeeds. This paper principally consists of two worries for Yli-Vakkuri and Hawthorne's argument and its relation to the formal system they develop. In my mind, these worries are simply that: worries. They are troubling enough to undermine confidence in this program's success—they do not ensure its failure. Nevertheless, much would need to be done to restore confidence in their result. The first problem I raise is that their assumptions entail the success of counterfactual strengthening—the inference from 'If *A* were true then *C* would be true' to 'If *A* and *B* were true then *C* would be true.' Many deny the felicity of counterfactual strengthening in ordinary modal contexts. Indeed, the Stalnaker (1968)/Lewis (1973*a*) semantics for counterfactual conditionals, which remains dominant in the discipline at large, entails that counterfactual strengthening fails. Whether Yli-Vakkuri and Hawthorne's assumptions are tenable depends (at least partially) on whether the problematic implications of strengthening can be derived in the language of mathematics. This requires a more precise account of what constitutes pure mathematics than is currently available. The second problem is that the assumptions which are responsible for this result are not those which Yli-Vakkuri and Hawthorne defend. Their result stems from adopting a type of entailment which validates the Necessitation Rule and Deduction Theorem, yet fails to distinguish an argument's premises from its axioms. Any such system has the very same result; it has nothing to do with mathematics in particular. Without a defense of this notion of entailment, we remain without a compelling argument for the necessity of mathematics.

Before turning to the details of Yli-Vakkuri and Hawthorne's account, a brief note on a tension between the worries I raise. While the first might reasonably be interpreted as the claim that their assumptions are far too strong (in that they force our hand on longstanding and seemingly intractable debate about counterfactual logic), the second is that the bulk of these very same assumptions are too weak to secure any theoretically interesting results (indeed, hardly more than are required to ensure that the language of mathematics is capable of expressing any modal claims at all). I will attempt to alleviate this tension in some concluding remarks; for the moment I simply note that it exists.

The Necessity of Mathematics

Yli-Vakkuri and Hawthorne's program fits broadly within a reorientation occurring in metaphysics. Following the formalization of modal logic in the 1960's, and the apparent theoretical uses for modality that ensued, many took possibility or necessity to be primitive, and defined other modal notions (such as the counterfactual conditional) in terms of them. In contrast, some contemporary philosophers maintain that the counterfactual conditional ought to be taken as primitive, and necessity and possibility defined in terms of it.[6] The crucial definition of necessity in terms of counterfactuality is the following:

$$\Box A =_{df} \neg A \mathrel{\Box\!\!\rightarrow} \bot$$

The claim that it is necessary that A amounts to the claim that if A were false, then the absurd would obtain. This definition receives support on several fronts. It is an immediate consequence of the aforementioned Stalnaker/Lewis semantics for counterfactual conditionals, according to which sentences of the form 'If A were true then B would be true' hold just in case the closest possible worlds in which A is true are also possible worlds in which B is true.[7] But perhaps the most compelling defense of this principle occurs in Williamson (2007), who demonstrates that it follows from a **K** modal logic—the weakest modal logic standardly available—and the following two principles:

[6]See Williamson (2007). This trend is in its infancy; it remains to be seen whether it will stand the test of time. Part of the motivation for this approach is that, Williamson maintains, we have more direct epistemic access to counterfactual conditionals than we have to necessity and possibility. While scientific experiments may inform us of what would happen if electrons were to pass through an open slit, it is not obvious that they inform us that water is necessarily H_2O. However, for alternate accounts of our epistemic access to modality, see, e.g., Hale (2003); Lowe (2012); Kment (2018).

[7]This is a rough gloss on their views, which differ in philosophically important ways. In particular, Stalnaker's similarity relation selects a unique, most similar w' for each possible world w, and determines the truth of counterfactuals by what occurs in it. Lewis, in contrast, evaluates counterfactuals by truth at the closest possible worlds (plural) and does not assume that there is a unique most-similar world. Each version has benefits over the other. For example, it is a consequence of Lewis's—but not Stalnaker's—view that the Counterfactual Excluded Middle ($A \mathrel{\Box\!\!\rightarrow} B \vee A \mathrel{\Box\!\!\rightarrow} \neg B$) fails. I take it that these debates, important though they are, have no bearing on the current project.

NECESSITY: $\Box(A \rightarrow B) \rightarrow (A \mathrel{\Box\!\!\rightarrow} B)$

POSSIBILITY: $(A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (\Diamond A \rightarrow \Diamond B)$

These assert, respectively, that if it is necessary that if A then B, then if A were to obtain then B would obtain, and that if it is the case that if A were to obtain then B would obtain, then if it is possible that A then it is possible that B. With the counterfactual definition of necessity at hand, possibility can be defined in the standard way:

$$\Diamond A =_{df} \neg\Box\neg A$$ [8]

With an eye toward the necessity of mathematics, Yli-Vakkuri and Hawthorne appeal to counterfactual conditionals occurring in mathematical texts. Sentences like "[If] there were a machine computing t [then] it would have some number k of states" (Boolos, Burgess and Jeffrey (2007)) regularly appear, and are naturally interpreted as counterfactual conditionals. Given that the truth-values of these sentences depend upon merely possible situations, mathematics is plausibly committed to a wide modal scope.

There is a natural objection to this interpretation which ought to be set aside. Arguably, counterfactual conditionals with necessary or impossible antecedents are somehow defective. A counterfactual with a necessary antecedent may collapse into the material conditional (because the closest world in which the antecedent obtains is the actual world), and a counterfactual with an impossible antecedent may be ill-formed (because there are no worlds in which the antecedent obtains).[9] Given the charitable assumption that mathematicians' assertions are neither trivial nor ill-formed, some might reasonably prefer alternate interpretations of Boolos, Burgess and Jeffrey's sorts of claims. However, it is worth recalling that Yli-Vakkuri and Hawthorne's imagined interlocutors are those who maintain that mathematical truths are contingent; they cannot object by appealing to the inadmissibility of counterfactuals with necessary or impossible antecedents, because they do not believe that mathematical counterfactuals *have* necessary or impossible antecedents.

Yli-Vakkuri and Hawthorne assume that the language of pure mathematics is at least equipped with sentences (which are denoted by 'A,' 'B,' etc. for individual sentences and by 'Γ,' Π,' etc. for collections of sentences), the classical logical connectives, the counterfactual connective $\mathrel{\Box\!\!\rightarrow}$, the absurdity operator $\perp$ and a symbol for informal provability $\vdash$. The least familiar of these is, presumably, the notion of informal provability.

[8]Those operating with an intuitionist modal logic would probably reject this definition of possibility. For the purposes of this paper, I follow Yli-Vakkuri and Hawthorne in assuming classical logic.

[9]This is the standard Stalnaker/Lewis line. There has, however, been a sustained defense of counterpossibles: counterfactual conditionals with impossible antecedents. See, for example, Cohen (1987); Mares (1997); Goodman (2004); Bjerring (2013); Brogaard and Salerno (2013). Nevertheless, I note that Yli-Vakkuri and Hawthorne do not avoid the collapse of the counterfactual conditional into the material conditional. As I mention below, it is provable on their assumptions that $A \mathrel{\Box\!\!\rightarrow} B$ iff $A \rightarrow B$.

Informal proofs are mathematically rigorous; the main difference between informal and formal proofs is that the results of informal proofs are universally true, while falsehoods are formally provable in systems with false axioms. Additionally, the notion of informal provability is sensitive to mathematical practice: the fact that mathematicians regularly license a particular kind of inference is evidence that it is admissible in informal proofs.

In addition to the counterfactual definition of necessity, Hawthorne and Yli-Vakkuri make the following assumptions:

CLASSICAL CONSEQUENCE	$\Gamma \vdash A$ whenever A follows from Γ by classical logic.
MODUS PONENS	$\Gamma, A \Rightarrow B, A \vdash B$ where $\Rightarrow$ is either the counterfactual or material conditional.
CUT	If $\Gamma \vdash A_1, ..., A_n$ and $\Pi, A_1, ..., A_n \vdash B$ then $\Pi, \Gamma \vdash B$.
COUNTERFACTUAL DEDUCTION	If $\Gamma, A \vdash B$, then $\Gamma \vdash A \mathrel{\Box\!\!\rightarrow} B$.
DEDUCTION THEOREM	If $\Gamma, A \vdash B$, then $\Gamma \vdash A \rightarrow B$.

Classical Consequence, Modus Ponens, Cut and Deduction Theorem are all, they claim, uncontroversial. The novel assumption is Counterfactual Deduction. But there is plenty of textual evidence that mathematicians assume that it is true. Take, for example:

> Let us designate the set of all such Gödel numbers by R, and let us suppose that R is recursively enumerable. Then, since $R \neq \varnothing$, there would exist a recursive function $f(n)$ whose range is R. (Davis, 1958, pg. 78)

Davis recognizes that, under the assumption that R is recursively enumerable, it is provable that there is a function whose range is R. What he concludes, then, is a counterfactual: if R were recursively enumerable, then there would be a function with R as its range. This is an instance of Counterfactual Deduction.

Or consider an elementary proof that there are infinitely many prime numbers. Suppose, for reductio, that there were finitely many primes. In this case, these primes would have a product n. The number $n + 1$ would not be evenly divisible by any prime number (except the number 1, depending on whether 1 regarded as prime), and would therefore be prime. However, $n + 1$ is not a factor of n, because it is larger than n. Therefore, n would not be the product of all primes, which contradicts the former claim that it is the product of all primes.

Several counterfactuals occurred in this proof. The relevant inference occurs from what is *provable* from the claim that there are finitely many primes to what *would occur* were there finitely many primes. This too is an instance of Counterfactual Deduction. Notably, the other principles Yli-Vakkuri and Hawthorne rely upon receive no sustained defense or discussion.

With such principles at hand, the derivation of the necessity of mathematics is as follows. Let A be an arbitrary sentence in the language of mathematics. From Classical Consequence, we have:

$$A, \neg A \vdash \bot$$

Counterfactual Deduction then entails:

$$A \vdash \neg A \;\Box\!\!\rightarrow \bot$$

The counterfactual definition of necessity then gives us:

$$A \vdash \Box A$$

Deduction Theorem then entails:

$$\varnothing \vdash A \rightarrow \Box A$$

This does not simply assert that if a sentence is true then it is necessarily true; it makes the stronger claim that it is *provable* that if A is true then it is necessarily true.[10]

Replacing A with $\Box A$ and $\Diamond A$ yields:

4: $\vdash \Box A \rightarrow \Box\Box A$

5: $\vdash \Diamond A \rightarrow \Box\Diamond A$

Classical Consequence, Deduction Theorem, Modus Ponens and Cut collectively imply that:

$$(\neg A \;\Box\!\!\rightarrow \bot) \rightarrow A$$

From the counterfactual definition of necessity, we then have:

[10]Note that this need not conflict with the incompleteness of various mathematical systems. There may be many sentences A in the language of pure mathematics such that A is true but $\vdash A$ is false. What this asserts is that, even in these cases, $\vdash A \rightarrow \Box A$ remains true.

T: $\Box A \rightarrow A$

Additionally, the **K** axioms of $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$ and $\vdash A \rightarrow \vdash \Box A$ are both theorems. These suffice to axiomatize **S5** modal logic. And so, Yli-Vakkuri and Hawthorne conclude, mathematics is committed not only to its own necessity, but to an **S5** system in particular. Far from being agnostic about its modal commitments, mathematics determines the system of modal logic which governs its theorems' results.

A Worry Concerning Counterfactual Strengthening

It is my hope that the previous (admittedly somewhat cursory) overview conveys both the structure and the initial appeal of Yli-Vakkuri and Hawthorne's argument. This argument is incontrovertibly valid, so any disagreement must emanate from challenging their assumptions—assumptions which strike me as *prima facie* plausible.

As it turns out, these seemingly innocuous assumptions have surprising implications. In particular, they entail that counterfactual conditional collapses into the material conditional; within the language of pure mathematics, '$A \rightarrow B$' holds just in case '$A \mathrel{\Box\!\!\rightarrow} B$' holds. The derivation of the collapse is as follows:

1. $A \rightarrow B, A \vdash B$ — *Modus Ponens*
2. $A \rightarrow B \vdash A \mathrel{\Box\!\!\rightarrow} B$ — *1, Counterfactual Deduction*
3. $\vdash (A \rightarrow B) \rightarrow (A \mathrel{\Box\!\!\rightarrow} B)$ — *2, Deduction Theorem*
4. $A \mathrel{\Box\!\!\rightarrow} B, A \vdash B$ — *Modus Ponens*
5. $A \mathrel{\Box\!\!\rightarrow} B \vdash A \rightarrow B$ — *4, Deduction Theorem*
6. $\vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \rightarrow B)$ — *5, Deduction Theorem*
7. $\vdash (A \rightarrow B) \leftrightarrow (A \mathrel{\Box\!\!\rightarrow} B)$ — *3, 6 Cut and Classical Consequence*

For example, it is provable that 'If $2+2 = 4$, then $2+3 = 5$' obtains if and only if 'If it were the case that $2 + 2 = 4$, then $2 + 3$ would equal 5' obtains. While this particular example is seemingly unproblematic, the collapse has undesirable implications. In particular, it forces our hand on a contentious debate between the following three principles of counterfactual logic:

SUBSTITUTION OF EQUIVALENTS	If A is logically equivalent to B, then if $A \boxright C$ then $B \boxright C$.
SIMPLIFICATION	If $(A \vee B) \boxright C$ then $A \boxright C$ and $B \boxright C$.
FAILURE OF COUNTERFACTUAL STRENGTHENING	It is not the case that $A \boxright C$ entails $(A \wedge B) \boxright C$.

Each of these principles has received some measure of support. The Substitution of Equivalents is often defended on theoretical grounds. If two sentences are logically equivalent, it is difficult to see how any difference between them could affect the truth-values of counterfactuals they occur within. After all, they hold in precisely the same possible situations. Additionally, it is an immediate consequence of the Stalnaker/Lewis semantics for counterfactual conditionals that the Substitution of Equivalents holds. The closest possible worlds in which a sentence obtains are invariably the closest possible worlds in which equivalent sentences obtain, so accounts that rely upon the closeness of worlds do not distinguish between equivalent expressions. Even when the commitment to a particular semantics for counterfactual conditionals is dropped, many endorse a principle allowing for the substitution of equivalent expressions.[11]

Simplification is often defended by appeal to ordinary reasoning.[12] It would be strange to assert 'If Jack or Jill were to come to the party then the party would be fun, and if Jack were to come to the party, it would not be fun.' Similarly, it seems reasonable for someone to deny 'If it were to rain or not to rain, then the street would be wet' on the grounds that they deny 'If it were not to rain, then the street would be wet.' Both of these examples involve appeals to Simplification.

Similarly, the Failure of Counterfactual Strengthening is often defended by appeal to the intuitive consistency of Sobel sequences.[13] It may be that if Tim were to take the aspirin, he would be fine, but if Tim were to take the aspirin and the cyanide, he would not be fine, and it may be that if the Federal Reserve were to lower the interest rate, the economy would grow, but if the Federal Reserve were to lower the interest rate and the European markets were to collapse, the economy would not grow. If these sentences are consistent, as they naturally seem to be, then counterfactual strengthening fails at least some of the time. Notably, this is a respect in which the counterfactual conditional appears to diverge from the material conditional. It is straightforward to establish that the material analog of

[11]For an extended discussion of how substitution coheres with natural-language modals, see Kratzer (1981*a*,*b*, 1986, 1991).

[12]This was independently noticed by Fine (1975) and Nute (1975) in response to Lewis (1973*a*). For a response to Nute, see Loewer (1976), and for the ensuing discussion about disjunctive antecedents in counterfactual conditionals more generally, see Lewis (1977); Nute (1980); Alonso-Ovalle (2006).

[13]See Sobel (1970). For canonical discussions of Sobel sequences, see Stalnaker (1968); Lewis (1973*a*,*b*).

Counterfactual Strengthening universally holds; that is, if $A \rightarrow C$ then $(A \wedge B) \rightarrow C$.

Despite these three principles' initial appeal, one must be abandoned, for they are mutually inconsistent. The conflict between them can be brought out in the following way:

1. $A \mathrel{\Box\!\!\rightarrow} C$ — *Supposition*
2. $A \vee (A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ — *1, Substitution of Equivalents*
3. $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ — *2, Simplification*

If Substitution of Equivalents and Simplification are both true, it follows that counterfactual strengthening universally succeeds. The two collectively entail that if 'If Sarah were to work hard, she would get a raise' is true, then 'If Sarah were to work hard and slap her boss, she would get a raise' is true as well.

While it is indisputable that these principles are incompatible, what we ought to do in light of this incompatibility is a matter of heated debate. Arguably, the most popular option is to retain the Substitution of Equivalents and the Failure of Counterfactual Strengthening, and to abandon Simplification. This option is forced upon us by the Stalnaker/Lewis semantics for counterfactuals. As mentioned before, this semantics licenses the Substitution of Equivalents, because equivalent expressions are true in the same possible situations. It also provides an intuitive explanation for the Failure of Counterfactual Strengthening. It may be that the closest worlds in which Sarah works hard are ones in which she gets a raise, but the closest worlds in which Sarah both works hard and slaps her boss are not ones in which she gets a raise, because the closest worlds in which she works hard are not ones in which she slaps her boss. Simplification fails when only one disjunct is relevant to the most-similar possible worlds. Perhaps all of the closest worlds in which either Jack or Jill come to the party are ones in which Jill comes to the party. In this case, the closest worlds in which Jack comes to the party are not relevant in determining the truth-value of 'If Jack or Jill were to come to the party, then the party would be fun.' Admittedly, abandoning Simplification is a theoretical cost, but the pertinent cases can arguably be accommodated pragmatically, rather than semantically.[14]

Others disagree. Recently, Fine (2012) provided a hyperintensional semantics for counterfactual conditionals—one which preserves both Simplification and the Failure of Counterfactual Strengthening and abandons the Substitution of Equivalents. Santorio (2018) advocates abandoning both the Substitution of Equivalents and Simplification, but preserves the Failure of Counterfactual Strengthening. And Kocurek (Forthcoming) provides independent reasons to abandon the Substitution of Equivalents. All counterpossibles (counterfactual conditionals with impossible antecedents) have equivalent antecedents,

[14]For pragmatic accounts of this phenomenon, see, e.g., Klinedinst (2009).

and few license the substitution of any impossible antecedent with another. If substitution principles fail for counterpossibles, it is reasonable to expect them to fail for ordinary counterfactuals as well. Debate rages on. While the Stalnaker/Lewis line remains prominent (minimally, given the enduring popularity of this semantics, it is an option many are tacitly committed to), it is safe to say that the fact that it forces our hand on this debate counts among its most controversial implications.

As it turns out, Yli-Vakkuri and Hawthorne's assumptions also force our hand in this debate, *but force it differently than Stalnaker and Lewis do*. Due to the collapse of the counterfactual conditional into the material conditional, their assumptions entail that the Substitution of Equivalents and Simplification are both true. Consequently, these assumptions entail that Counterfactual Strengthening universally succeeds.

The derivation of the Substitution of Equivalents is as follows:

Suppose that A is logically equivalent to B.

1. $B \vdash A$ — *Classical Consequence*
2. $A \mathrel{\Box\!\!\rightarrow} C, A \vdash C$ — *Modus Ponens*
3. $A \mathrel{\Box\!\!\rightarrow} C, B \vdash C$ — *1, 2 and Cut*
4. $A \mathrel{\Box\!\!\rightarrow} C \vdash B \mathrel{\Box\!\!\rightarrow} C$ — *3, Counterfactual Deduction*
5. $\vdash (A \mathrel{\Box\!\!\rightarrow} C) \rightarrow (B \mathrel{\Box\!\!\rightarrow} C)$ — *4, Deduction Theorem*

The derivation of Simplification is as follows:

6. $A \vdash A \vee B$ — *Classical Consequence*
7. $(A \vee B) \mathrel{\Box\!\!\rightarrow} C, A \vee B \vdash C$ — *Modus Ponens*
8. $(A \vee B) \mathrel{\Box\!\!\rightarrow} C, A \vdash C$ — *6, 7 and Cut*
9. $(A \vee B) \mathrel{\Box\!\!\rightarrow} C, \vdash A \mathrel{\Box\!\!\rightarrow} C$ — *8, Counterfactual Deduction*
10. $\vdash ((A \vee B) \mathrel{\Box\!\!\rightarrow} C) \rightarrow (A \mathrel{\Box\!\!\rightarrow} C)$ — *9, Deduction Theorem*

As with the derivation of the necessity of mathematics, it not only follows that the Substitution of Equivalents and Simplification hold, but it is always provable that they hold. As we have already seen, these principles, in turn, entail the success of Counterfactual Strengthening. Therefore, Yli-Vakkuri and Hawthorne's assumptions entail that every instance of Counterfactual Strengthening expressible in the language of pure mathematics succeeds.

How troubling this result? Presumably, this depends (at least partially) on what can be expressed within the language of pure mathematics. If Sobel sequences are expressible, then the assumptions have untenable implications; few react to the conflict between Substitution, Simplification and Strengthening by jettisoning Strengthening—the plausible consistency of Sobel Sequences seems indispensable to modal reasoning. Determining the viability of this program thus requires an account of what constitutes pure mathematics. Without one, it is challenging to determine whether Sobel Sequences can be expressed. It would be question-begging to identify pure mathematics with those branches that are necessarily true—some other characterization is required.

Yli-Vakkuri and Hawthorne do not specify what the boundaries of pure mathematics are. It is partially for this reason that my concern is merely a worry, rather than a knock-down critique. Perhaps their view could be supplemented by an account of pure mathematics—one which evades the problems that Counterfactual Strengthening generates. However, there is reason to suspect that they face such a worry: that the language they are concerned with has the resources to express Sobel Sequences.

One place this crops up is in response to a potential response to an objection they consider—the dispensability objection. Some might suspect that mathematical counterfactuals are dispensable to mathematics. Mathematicians may employ them in order to improve readability or to add linguistic flair, but they could be removed without affecting any substantive result. If mathematical counterfactuals are dispensable, we ought not derive substantive modal conclusions from them. This worry is compounded by the collapse from the counterfactual to the material conditional; in every case where mathematicians employ a counterfactual conditional, they could employ the material conditional instead.

Yli-Vakkuri and Hawthorne deny that mathematical counterfactuals are dispensable, claiming the following:

> Counterfactuals are absolutely indispensable to what mathematics contributes to our total body of knowledge...Note first that myriad applications of mathematics to the hustle and bustle of both everyday life and engineering require our knowing that mathematical truths would remain true even if things had gone differently in various ways. For example, in justifying a particular engineering solution, one often appeals to mathematical truths in reasoning about how things would have gone if one had opted for an alternative solution. In doing so one assumes—and if one is successful, one knows—that those mathematical truths would have been true even if one had opted for the alternative solution. Note second that, as the queen of the sciences, mathematics is primed for application in any area of objective inquiry, whether it be the science of electromagnetism, the theory of rook and pawn endings, or natural language semantics. (Pg. 14)

This passage strongly suggests that mathematical counterfactuals occur in disciplines ranging from engineering to electromagnetism to natural language semantics. After all, if the language of mathematics is incapable of expressing these counterfactuals, how could they lend support for the indispensability of counterfactuals in mathematics? Sobel Sequences are derivable in every discipline they mention. An engineer might derive the fact that, if a pulley were to double in size, it could lift a heavy box, but would deny that if a pulley were to double in size and be made of twine, it could lift a heavy box. A physicist might conclude that if an electron were to be placed in a field, it would accelerate, but deny that if an electron were to be placed in a field and an equal-but-opposite force were to be introduced, it would accelerate. Both the engineer and physicist thus deny the felicity of Counterfactual Strengthening in the counterfactuals they appeal to. Yli-Vakkuri and Hawthorne's assumptions, which entail that counterfactual strengthening succeeds, are at odds with this practice.

Of course, it might be claimed that sentences occurring in engineering, physics, and the like are not pure mathematics. Yli-Vakkuri and Hawthorne's argument for the indispensability of mathematical counterfactuals is arguably an appeal to *applied* mathematics—not to pure mathematics. As such, these sentences fall outside of the scope of their program; they need not claim that Counterfactual Strengthening is admissible in these types of cases, because these sentences are not within the language of pure mathematics. Of course, once it is claimed that these sentences do not count as purely mathematical, the dispensability objection returns (after all, it may be that counterfactuals are indispensable to applied mathematics, but are they indispensable to pure mathematics?). But that is not the problem I am presently concerned with. Are there other—purer—cases where Counterfactual Strengthening arises?

Some purely mathematical cases seem innocuous. It seems reasonable to accept that 'If 2 were prime, there would be an even prime' and 'If 2 were prime and 3 were prime, there would be an even prime' are both perfectly true. However other cases are much more suspect. Consider, for example, the simple arithmetic statement 'If 6 were added to 7, the result would be 13.' This sentence does not entail 'If 6 were added to 7 and 5 were subtracted, the result would be 13.' After all, the result would be 8, not 13. And yet the second sentence is the strengthened version of the first; if counterfactual strengthening holds, the latter ought to be true if the former is. Consider, also, the relation between 'If there were a Turing machine in state T, then two steps later it would be in state T'' and 'If there were a Turing machine in state T and a 0 were changed to a 1, then two steps later it would be in state T'.' The first of these sentences may be true while the second may be false. Once again, however, the second sentence is a strengthened version of the first, so if counterfactual strengthening holds then the truth of the first ought to guarantee the truth of the second.

The same maneuver is available here as was available for the dispensability objection. It might be maintained that these sentences fall outside the purview of pure mathematics, and so are not expressible in the language Yli-Vakkuri and Hawthorne have in mind. However, it is difficult to see why this would be the case. By stipulation, pure mathematics is capable of expressing the counterfactual conditional. The only other terms (in the first example) concern numbers, primeness, evenness, and the results of arithmetic calculations. What constitutes pure mathematics, if not this?

It thus seems that there are numerous examples where mathematical expressions do not permit counterfactual strengthening. But without an account of pure mathematics, Yli-Vakkuri and Hawthorne may always retreat; they may claim that the examples I discuss cannot be expressed within the language of pure mathematics, and, as such, do not undermine their account. But with each retreat, their project becomes more limited in scope. Their project does not secure the necessity of sentences which fall outside the scope of pure mathematics.

When I first considered this problem it seemed to me there was an additional interpretive puzzle for Yli-Vakkuri and Hawthorne: it is not at all obvious what the semantics underlying mathematical counterfactuals is. The Stalnaker/Lewis semantics requires Simplification to be false, while Yli-Vakkuri and Hawthorne's assumptions entail that Simplification is true. The Stalnaker/Lewis approach remains the dominant interpretation of counterfactuals in philosophy (and beyond). Without the ability to appeal to it, another ought to take its place. Absent any semantics for mathematical counterfactuals at all, it is unclear what, precisely, they mean.

However, there is an interpretation of these counterfactuals that has the logical attributes Yli-Vakkuri and Hawthorne desire. This interpretation is sometimes referred to as a 'strict counterfactual implication,' and predates the Stalnaker/Lewis semantics.[15] On this view, the counterfactual conditional $A \mathrel{\Box\!\!\rightarrow} B$ is synonymous with $\Box(A \rightarrow B)$. The claim that if A were true then B would be true amounts to the claim that it is necessary that if A is true then B is true. This interpretation can then be supplemented by the standard Kripke semantics for necessity and possibility (or any other such semantics) to furnish a semantics for counterfactual conditionals. The reason this interpretation is amenable to Yli-Vakkuri and Hawthorne's program arises firstly from the fact that the counterfactual conditional collapses into the material conditional (i.e., that $(A \mathrel{\Box\!\!\rightarrow} B)$ iff $(A \rightarrow B)$), and, secondly, from the fact that every truth within their language is necessarily true (i.e., that $(A \rightarrow B)$ iff $\Box(A \rightarrow B)$).

Although this approach renders mathematical counterfactuals meaningful, it does so at a cost. The semantics underlying mathematical counterfactuals differs from the semantics of ordinary counterfactual conditionals. Few, if any, contemporary logicians endorse the strict interpretation of counterfactual conditionals in non-mathematical contexts, precisely because the strict interpretation licenses counterfactual strengthening. The strict interpretation entails that 'If Julia were to take the bus, she would save money' implies that 'If Julia were to take the bus and buy a Ferrari, she would save money.' After all, in order for the first sentence to be true, it must be necessary that if Julia takes the bus, then she

[15]This type of view was endorsed by, for example, Pierce (1896). How extensive its history is is a matter of debate. Pierce attributes this sort of view to the Hellenistic logician Philo. However, Bobzein (2011) prefers an alternate intrepretation of Philo's work.

saves money. This approach thus requires the semantics for mathematical counterfactuals to come apart from the semantics of ordinary counterfactuals—the two types of expressions mean different things. What's more, this is not peculiar to the strict interpretation of counterfactual conditionals. Because Yli-Vakkuri and Hawthorne's assumptions entail that the logic of mathematical counterfactuals differs from the logic of ordinary counterfactuals, any semantics for mathematical counterfactuals must likewise diverge from a semantics for ordinary counterfactuals.[16] So while it may be possible for Yli-Vakkuri and Hawthorne to supplement their view with a semantics, any they appeal to will require that counterfactuals occurring in mathematical contexts mean something different from counterfactuals occurring in ordinary contexts.

A Worry Concerning Counterfactual Metalogic

My second worry could be framed in multiple ways, perhaps the most charitable of which is this: Yli-Vakkuri and Hawthorne defend the wrong thing. The only assumption which receives a substantive defense is Counterfactual Deduction. This, we may recall, is defended by appeal to mathematical practice. All other assumptions are treated as orthodox principles of counterfactual logic. However, I maintain that, rather than Counterfactual Deduction, the assumptions (one of which is implicit) which are entirely responsible for their result are the following:

i) The Deduction Theorem: If $A \vdash B$ then $\vdash A \to B$.

ii) The Necessitation Rule: If $\vdash A$ then $\vdash \Box A$.

iii) The claim that $\vdash$ does not distinguish between an argument functioning as a premise and as an axiom.

To see what these assumptions entail, let us select an arbitrary sentence A—for the sake of intelligibility, let A be the sentence 'Grass is green.' We wish to investigate what it is that A entails, and so we include A as a premise in an argument. Given assumption iii), this effectively adds A to the set of axioms within our system. With the Necessitation Rule in place, we may conclude that every axiom—including A—is necessarily true. After all, axioms are theorems in this proof-theoretic system, and the Necessitation Rule allows us to conclude that theorems are necessarily true. That is to say, given both the necessitation rule and the failure to distinguish between premises and axioms, we may derive $A \vdash \Box A$. With Deduction Theorem also in place, one may then infer $\vdash A \to \Box A$: the claim that if

[16] More precisely, any semantics for which their assumptions are both sound and complete will not be a semantics for which the logic of ordinary counterfactual conditionals is sound and complete—at least on the assumption that Counterfactual Strengthening fails for ordinary counterfactuals.

grass is green then it is necessary that grass is green.

This result should seem eerily familiar. It is of precisely the same form as Yli-Vakkuri and Hawthorne's conclusion regarding the necessity of mathematics. And, as with their conclusion, it not only follows that if A is true then A is necessarily true, but that it is *provable* that if A is true then it is necessarily true. However, this result is not peculiar to mathematics: any system which licenses Deduction Theorem and the Necessitation Rule while failing to distinguish an argument's premises from its axioms has this result.

In many contexts, of course, this result is entirely implausible; it would be absurd to claim that if grass is green then it is necessary that grass is green. Quite generally, we are rarely willing to accept that if a sentence is true then it is necessarily true. It is precisely for this reason that modal logicians typically rejected (at least one of) assumptions i-iii.

For example, numerous authors claim that Deduction Theorem fails for modal logic.[17] Deduction Theorem is extremely intuitive, and is an immediate metatheorem of propositional and first-order logic.[18] It holds in many nonclassical systems as well. However, precisely to avoid the conclusion that $\vdash A \to \Box A$, many have rejected it for modal logic. For example, Fitting (2007) states:

> "Modal logic raises problems for the notion of deduction. Suppose we want to show $X \to Y$ in some modal axiom system by deriving Y from X. So we add X to our axioms. Say, to make things both concrete and intuitive, that X is 'it is raining' and Y is 'it is necessarily raining.' Since X has been added to the axiom list the necessitation rule applies, and from X we conclude $\Box X$, that is Y. Then Deduction Theorem would allow us to conclude that if it is raining, it is necessarily raining. This does not seem right—nothing would ever be contingent."

Fitting is hardly an outlier in this regard. Others who reject Deduction Theorem for modal logic include (but are not limited to) Smorynski (1984); Fagin et al. (1995); Chagrov and Zakharyaschev (1997); Ganguli and Nerode (2004); Sider (2010). If Deduction Theorem fails for modal logic, then the conclusion that $\vdash A \to \Box A$ can be avoided.

Necessitation can be blocked in other ways. Recently, Hakli and Negri (2012) argue that Deduction Theorem succeeds for modal logic. They avoid necessitation, instead, by rejecting assumption iii—the claim that $\vdash$ does not distinguish between an argument functioning as a premise and as an axiom, and further distinguish two notions of entailment which result: one in terms of truth and another in terms of validity. The distinction

[17]Debates about the viability of Deduction Theorem for modal logic have occurred since the its formalism. See Barcan (1946); Barcan Marcus (1953); Feys (1965) for examples of early philosophers who dispute Deduction Theorem for at least some modal systems. These early debates predominantly concern different problems than the one I primarily address.

[18]Kleene (1952) states that Deduction Theorem was first proved for propositional and first-order logic by Herbrand (1930 (1971)).

between these two types of entailment was noted independently by Avron (1991) in connection to first-order logic, who defines it as follows:

TRUTH: $\Gamma \vdash_t A$ iff every assignment in a first-order structure which makes Γ true also makes A true.

VALIDITY: $\Gamma \vdash_v A$ iff if Γ is valid (i.e., true on all assignments), then A is true.

Notably, while $A(x) \vdash_v \forall x F(x)$, $A(x) \nvdash_t \forall x F(x)$. On some assignments $A(x)$ is true while $\forall x F(x)$ is false, but if it is the case that $A(x)$ is true on all assignments, then $\forall x F(x)$ is as well. Halki and Negri offer a similar distinction for modal logic:

TRUTH: $\Gamma \vdash_t A$ iff given a frame and a valuation in that frame and a world in it, if Γ are true in that world then A is true in that world.

VALIDITY: $\Gamma \vdash_v A$ iff given a frame if Γ are true in every world, then A is true.

It should be no surprise that $\Box$ functions analogously to $\forall$—the terms act similarly in many logical respects. In particular, a parallel result to Avron's holds: while $A \vdash_v \Box A$, $A \nvdash_t \Box A$. After all, if A is true in every world it follows that $\Box A$ is true as well, but if A is only true at a particular world in a frame, it does not follow that $\Box A$ is true at that world in a frame, as there may be accessible worlds in which A is false. So the necessitation rule (when applied to a premise, rather than an axiom) holds for $\vdash_v$ but not for $\vdash_t$. In contrast, while $A \vdash_t B$ entails $\vdash_t A \rightarrow B$, but $A \vdash_v B$ does not entail $\vdash_v A \rightarrow B$. So while Deduction Theorem succeeds for $\vdash_t$ it fails for $\vdash_v$. There are thus two notions of entailment on Hakli and Negri's system, one of which validates the Necessitation Rule, and the other of which validates Deduction Theorem, but neither of which validates both the Necessitation Rule and Deduction Theorem. And so, by distinguishing between an argument functioning as a premise and as an axiom, they avoid necessitation.

At the outset of this section, I charitably framed my worry as a need to redirect support to claims i)-iii). A less charitable interpretation is this: Yli-Vakkuri and Hawthorne have simply rediscovered that these three assumptions entail $\vdash A \rightarrow \Box A$. This result has nothing whatsoever to do with mathematics. What is special about mathematics is only the fact that the conclusion is plausible in this case.

Why believe Yli-Vakkuri and Hawthorne have merely rediscovered this result? Because their assumptions are provably equivalent to appending to appending Deduction Theorem to a **T** modal logic in a system which fails to distinguish between premises and axioms.

I establish this indirectly. What I immediately prove is that their system is equivalent to combining Deduction Theory to a system of counterfactual logic formalized by Williamson

(2007).[19] Each set of axioms can be used to derive the other. Independently, Williamson proves that his system is equivalent to **T**; it is an immediate consequence that their logic is equivalent to appending Deduction Theorem to **T**. Williamson's assumptions are the following:

PC	If A is a truth-functional tautology then $\vdash A$
REFLEXIVITY	$\vdash A \mathrel{\Box\!\!\rightarrow} A$
VACUITY	$\vdash (\neg A \mathrel{\Box\!\!\rightarrow} A) \rightarrow (B \mathrel{\Box\!\!\rightarrow} A)$
MP	If $\vdash A \rightarrow B$ and $\vdash A$ then $\vdash B$
MP$\Box$	$\vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \rightarrow B)$
CLOSURE	If $\vdash B \rightarrow C$ then $\vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \mathrel{\Box\!\!\rightarrow} C)$
EQUIVALENCE	If A is equivalent to $A*$ then $\vdash A \mathrel{\Box\!\!\rightarrow} B$ iff $\vdash A* \mathrel{\Box\!\!\rightarrow} B$

MP$\Box$ is sometimes referred to as 'weak strengthening'—it corresponds to the **T** axiom $\Box A \rightarrow A$: the axiom that if a claim is necessary then it actually holds (which, in turn, corresponds to the assumption that accessibility is a reflexive relation). Williamson also proves that fragment of this system without *MP*$\Box$ is equivalent to **K**—the weakest modal logic standardly available, which is characterized by the Necessitation Rule and the Distributive Axiom $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$.

Recall that Yli-Vakkuri and Hawthorne's assumptions included two versions of Modus Ponens: one in terms of the material conditional and the other in terms of the counterfactual conditional. If we consider the fragment of their system without the counterfactual version of Modus Ponens, their system is provably equivalent to Williamson's system without *MP*$\Box$, (and, therefore to **K**), when appended to Deduction Theorem. Notably, the counterfactual version of Modus Ponens plays no role in their derivation of the necessity of mathematics. This arises, I maintain, because the **T** axiom is gratuitous in deriving $\vdash A \rightarrow \Box A$; all that matters is necessitation, deduction theorem, and the appropriate notion of entailment. However, the **T** axiom (and, correspondingly, the counterfactual version of Modus Ponens) is indispensable in deriving the commitment to **S5**. If we are to demonstrate that mathematical practice is committed to an **S5** modality, it must be shown that mathematical practice is committed to the counterfactual version of Modus Ponens.

The upshot is this: Yli-Vakkuri and Hawthorne's assumptions are provably equivalent to appending Deduction Theorem to a **T** modal logic in a system which treats premises as

[19]See Appendix for the details of this proof.

axioms. It has already been established that such a system results in necessitation. In most contexts, this result is untenable, and leads logicians to reject assumptions that this result turns upon. What is special about mathematics is merely the fact that this conclusion appears plausible in this case. In order for Yli-Vakkuri and Hawthorne's result to secure the foundations for the necessity of mathematics, it must be shown that the necessitation rule holds, that Deduction Theorem is admissible, and that informal provability does not distinguish between premises and axioms. We presently have no defense of Deduction Theorem and the indiscriminate notion of informal provability, and so we lack a basis for the necessity of mathematics.

Conclusion

At the outset, I noted a tension between the two worries I raise: that, while the first could be understand as the claim that Yli-Vakkuri and Hawthorne's assumptions have implausible implications about counterfactual logic, the second is that the bulk of these assumptions perform minimal theoretical work. I do not believe these concerns are at odds. This is because while many of their assumptions are innocuous, Deduction Theorem and the claim that provability does not distinguish premises from axioms are not. It is no surprise that Deduction Theorem arises in both the derivations of the Substitution of Equivalents and Simplification; it carries weight both in deriving the necessity of mathematics as well as the controversial implications that Yli-Vakkuri and Hawthorne's theory has.

These worries I have raised are inconclusive—it may yet be that the challenges can be answered. But what is needed is a substantive defense of principles where now there is none.

References

Alonso-Ovalle, Luis. 2006. "Counterfactuals, Correlatives and Disjunction." *Linguistics and Philosophy* 32:207–44.

Avron, Arnon. 1991. "Simple Consequence Relations." *Information and Computation* 92.

Barcan Marcus, Ruth. 1953. "Strict Implication, Deducibility and Deduction Theorem." *The Journal of Symbolic Logic* 18:119–60.

Barcan, Ruth. 1946. "The Deduction Theorem in a Functional Calculus of First Order Based on Strict Implication." *The Journal of Symbolic Logic* 11:115–8.

Bealer, George. 2002. Modal Epistemology and the Rationalist Renaissance. In *Conceivability and Possibility*, ed. Tamar Gendler and Johnathan Hawthorne. Oxford University Press pp. 71–126.

Bjerring, Jens Christian. 2013. "On Counterpossibles." *Philosophical Studies* 190:327–53.

Bobzein, Susanne. 2011. "Dialectical School." *Stanford Encyclopedia of Philosophy* .

Boolos, George, John Burgess and Richard Jeffrey. 2007. *Computability and Logic*. Cambridge University Press.

Brogaard, Berit and Joe Salerno. 2013. "Remarks on Counterpossibles." *Synthese* 190:639–60.

Chagrov, Alexander and Michael Zakharyaschev. 1997. *Modal Logic*. Oxford University Press.

Cohen, Daniel. 1987. "The Problem of Counterpossibles." *Notre Dame Journal of Formal Logic* 29(9):91–101.

Davis, Martin. 1958. *Computability and Unsolvability*. McGraw-Hill.

Fagin, R., J Halpern, Y Moses and M Vardi. 1995. *Reasoning about Knowledge*. MIT Press.

Feys, Robert. 1965. Modal Logics. In *Collection de Logique Mathématique, Série B, IV*, ed. J Dopp. E. Nauwelaerts Éditeur.

Fine, Kit. 1975. "Review of Lewis's Counterfactuals." *Mind* 84:451–8.

Fine, Kit. 2012. "Counterfactuals Without Possible Worlds." *Journal of Philosophy* 109(3):221–46.

Fitting, Mel. 2007. Modal Proof Theory. In *Handbook of Modal Logic*, ed. P. Blackburn, J. can Benthem and F. Wolter. Elsevier pp. 85–138.

Frege, Gottlob. 1884. *The Foundations of Arithmetic*. Oxford University Press.

Ganguli, Suman and Anil Nerode. 2004. "Effective Completeness Theorems for Modal Logic." *Annals of Pure and Applied Logic* 128(1-3):141–95.

Gendler, Tamar and Jonathan Hawthorne. 2002. *Conceivability and Possibility*. Oxford University Press.

Goodman, Jeffrey. 2004. "An Extended Lewis/Stalnaker Semantics and the New Problem of Counterpossibles." *Philosophical Papers* 33:35–66.

Hakli, Raul and Sara Negri. 2012. "Does Deduction Theorem Fail for Modal Logic?" *Synthese* 187:849–67.

Hale, Robert. 2003. "Knowledge of Possibility and Necessity." *Proceedings of the Aristotelian Society* 103:1–20.

Hale, Robert and Crispin Wright. 2001. *The Reason's Proper Study: Essays Toward a Neo-Fregean Philosophy of Mathematics*. Oxford University Press.

Herbrand, Jacques. 1930 (1971). Recherches sur la theorie de la demonstration. In *Jacques Herbrand: Logical Writings*, ed. Warren Goldfarb. Harvard University Press.

Hodges, Wilfred. Forthcoming. "Necessity in Mathematics.".

Kleene, Stephan. 1952. *Introduction to Mathematics*. Van Nostrand.

Klinedinst, Nathan. 2009. "(Simplification of) Disjunctive Antecedents." *MIT Working Papers in Linguistics* 60.

Kment, Boris. 2018. "Essence and Modal Knowledge." *Synthese* pp. 1–23.

Kocurek, Alex. Forthcoming. "On the Substitution of Identicals in Counterfactual Reasoning." *Noûs* .

Kratzer, Angelika. 1981*a*. The Notional Category of Modality. In *Words, Worlds and Contexts: New Approaches to World Semantics*, ed. H. J. Eikmeyer and H. Rieser. de Gruyter.

Kratzer, Angelika. 1981*b*. "Partition and Reviesion: The Semantics of Counterfactuals." *Journal of Philosophical Logic* 10(2):201–16.

Kratzer, Angelika. 1986. "Conditionals." *Chicago Linguistics Society: Papers from the Parasession on Pragmatics and Grammatical Theory* 22(2):1–15.

Kratzer, Angelika. 1991. "Modality." *Semantics: an International Handbook of Contemporary Research* pp. 639–50.

Kripke, Saul A. 1980. *Naming and Necessity*. Harvard University Press.

Lewis, David. 1973*a*. *Counterfactuals*. Blackwell Publishers.

Lewis, David. 1973*b*. "Counterfactuals and Comparative Probability." *Journal of Philosophical Logic* 2(4):418–46.

Lewis, David. 1977. "Possible-World Semantics for Counterfactual Logics: A Rejoinder." *Journal of Philosophical Logic* 6(1):359–63.

Loewer, Barry. 1976. "Counterfactuals with Disjunctive Antecedents." *Journal of Philosophy* 73(16):531–7.

Lowe, E. J. 2012. "What is the Source of our Knowledge of Modal Truths?" *Mind* 121(484):919–50.

Mares, Edward. 1997. "Who's Afraid of Impossible Worlds?" *Notre Dame Journal of Formal Logic* 38:535–72.

Nute, Donald. 1975. "Counterfactuals and the Similarity of Worlds." *Journal of Philosophy* 72(21):773–8.

Nute, Donald. 1980. "Conversational Scorekeeping and Conditionals." *Journal of Philosophical Logic* 9(2):153–66.

Pierce, Charles. 1896. "Regenerated Logic." *The Monist* 7(1):19–40.

Santorio, Paolo. 2018. "Alternatives and Truthmakers in Conditional Semantics." *Journal of Philosophy* 115(10):513–49.

Sider, Theodore. 2010. *Logic for Philosophy*. Oxford University Press.

Smorynski, Craig. 1984. Modal Logic and Self-Reference. In *Handbook of Philosophical Logic, Volume II*, ed. D. Gabbay and F Guenthner. Kluwer.

Sobel, Howard. 1970. "Utilitarianisms: Simple and General." *Inquiry* 13(4):2008.

Stalnaker, Robert. 1968. A Theory of Conditionals. In *Studies in Logical Theory*, ed. N. Recher. Oxford University Press.

Williamson, Timothy. 2007. *The Philosophy of Philosophy*. Blackwell Publishing.

Yli-Vakkuri, Juani and John Hawthorne. 2018. "The Necessity of Mathematics." *Noûs* pp. 1–28.

Appendix

The following is a proof that the system of counterfactual logic developed by Yli-Vakkuri and Hawthorne (2018) is equivalent to appending Deduction Theorem to a **T** modal logic. This proof proceeds indirectly. What I immediately establish is that the system is equivalent to a counterfactual logic developed in Williamson (2007), when appended to Deduction Theorem. However, Williamson independently proves that his logic (without Deduction Theorem) is equivalent to **T**; it follows that Yli-Vakkuri and Hawthorne's system is equivalent to the conjunction of **T** with Deduction Theorem. The only additional assumption I make about Williamson's logic is that it is monotonic; i.e., that if $\Gamma \vdash B$ then $\Gamma, A \vdash B$.

I begin by establishing that Williamson's axioms follow from Yli-Vakkuri and Hawthorne's.

<u>PC:</u> If A is a truth-functional tautology, then $\vdash A$

This follows immediately from Classical Consequence; i.e., if A is a truth-functional tautology, then

$$\varnothing \vdash A \tag{1}$$

<u>REFLEXIVITY:</u> $\vdash A \mathrel{\Box\!\!\rightarrow} A$

Classical Consequence entails:

$$A \vdash A \tag{2}$$

(2) and Counterfactual Deduction then entail:

$$\varnothing \vdash A \mathrel{\Box\!\!\rightarrow} A \tag{3}$$

<u>VACUITY:</u> $\vdash (\neg A \mathrel{\Box\!\!\rightarrow} A) \rightarrow (B \mathrel{\Box\!\!\rightarrow} A)$

An instance of Modus Ponens is:

$$\neg A, \neg A \mathrel{\Box\!\!\rightarrow} A, B \vdash A \tag{4}$$

(4) and Deduction Theorem then entail:

$$\neg A \mathrel{\Box\!\!\rightarrow} A, B \vdash \neg A \rightarrow A \tag{5}$$

Classical Consequence entails:

$$\neg A \to A \vdash A \tag{6}$$

(5), (6), and Cut collectively entail:

$$\neg A \mathrel{\Box\!\!\rightarrow} A, B \vdash A \tag{7}$$

(7) and Counterfactual Deduction then entail:

$$\neg A \mathrel{\Box\!\!\rightarrow} A \vdash B \mathrel{\Box\!\!\rightarrow} A \tag{8}$$

And, finally, (8) and Deduction Theorem entail:

$$\varnothing \vdash (\neg A \mathrel{\Box\!\!\rightarrow} A) \to (B \mathrel{\Box\!\!\rightarrow} A) \tag{9}$$

MP: If $\vdash A \to B$ and $\vdash A$, then $\vdash B$.

Let us suppose the following:

$$\varnothing \vdash A \to B \tag{10}$$

$$\varnothing \vdash A \tag{11}$$

An instance of Modus Ponens—which is not to be confused with MP—is the following:

$$A \to B, A \vdash B \tag{12}$$

(10), (12) and Cut entail:

$$A \vdash B \tag{13}$$

(11), (13) and Cut then entail:

$$\varnothing \vdash B \tag{14}$$

MP□ (Weak Centering): $\vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \rightarrow B)$

An instance of Modus Ponens is the following:

$$A \mathrel{\Box\!\!\rightarrow} B, A \vdash B \tag{15}$$

(15) and Deduction Theorem entail:

$$A \mathrel{\Box\!\!\rightarrow} B \vdash A \rightarrow B \tag{16}$$

(16) and Deduction Theorem entail:

$$\varnothing \vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \rightarrow B) \tag{17}$$

CLOSURE: If $\vdash B \rightarrow C$ then $\vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \mathrel{\Box\!\!\rightarrow} C)$

Let us suppose that:

$$\varnothing \vdash B \rightarrow C \tag{18}$$

An instance of Modus Ponens is the following:

$$B, B \rightarrow C \vdash C \tag{19}$$

(18), (19) and Cut entail:

$$B \vdash C \tag{20}$$

Another instance of Modus Ponens is:

$$A, A \mathrel{\Box\!\!\rightarrow} B \vdash B \tag{21}$$

(20), (21) and Cut then entail:

$$A, A \mathrel{\Box\!\!\rightarrow} B \vdash C \tag{22}$$

(22) and Counterfactual Deduction then entail:

$$A \mathrel{\Box\!\!\rightarrow} B \vdash A \mathrel{\Box\!\!\rightarrow} C \tag{23}$$

And, finally, (23) and Deduction Theorem entail:

$$\varnothing \vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \mathrel{\Box\!\!\rightarrow} C) \tag{24}$$

<u>EQUIVALENCE:</u> If A is equivalent to $A*$, then $\vdash A \mathrel{\Box\!\!\rightarrow} B$ iff $\vdash A* \mathrel{\Box\!\!\rightarrow} B$

Let us suppose that:

$$A \leftrightarrow A* \tag{25}$$

I begin by establishing that if $\vdash A \mathrel{\Box\!\!\rightarrow} B$, then $\vdash A* \mathrel{\Box\!\!\rightarrow} B$. Let us suppose that:

$$\varnothing \vdash A \mathrel{\Box\!\!\rightarrow} B \tag{26}$$

(26) and the Monotonicity entail:

$$A* \vdash A \mathrel{\Box\!\!\rightarrow} B \tag{27}$$

An instance of Modus Ponens is:

$$A*, A, A \mathrel{\Box\!\!\rightarrow} B \vdash B \tag{28}$$

(27), (28) and Cut entail:

$$A*, A \vdash B \tag{29}$$

(25) and Classical Consequence entail:

$$A* \vdash A \tag{30}$$

(29), (30) and Cut entail:

$$A* \vdash B \tag{31}$$

(31) and Counterfactual Deduction then entail:

$$\varnothing \vdash A* \mathrel{\Box\!\!\rightarrow} B \tag{32}$$

A parallel proof establishes that if $\vdash A* \mathrel{\Box\!\!\rightarrow} B$ then $\vdash A \mathrel{\Box\!\!\rightarrow} B$. From this it follows that:

$$\vdash A \mathrel{\Box\!\!\rightarrow} B \text{ iff } \vdash A* \mathrel{\Box\!\!\rightarrow} B \tag{33}$$

<u>DEDUCTION THEOREM</u>

Deduction Theorem is an axiom in both systems under consideration; it follows trivially from itself.

Therefore, all of Williamson's axioms (with Deduction Theorem) follow from Yli-Vakkuri and Hawthorne's. In order to establish the equivalence of these systems, it is sufficient to prove Yli-Vakkuri and Hawthorne's axioms from Williamson's (with Deduction Theorem). I precede slightly out of order from Williamson's presentation in order to use earlier proofs to facilitate later ones. For the purposes of this paper, it suffices to demonstrate the unary instance of Cut (which, incidentally, is the only instance employed in their demonstration of the necessity of mathematics).

<u>CUT:</u> If $\Gamma \vdash A$ and $\Pi, A \vdash B$ then $\Gamma, \Pi \vdash B$.

Let us suppose that:

$$\Gamma \vdash A \tag{34}$$

$$\Pi, A \vdash B \tag{35}$$

(35) and Monotonicity entail:

$$\Gamma, \Pi, A \vdash B \tag{36}$$

(36) and Deduction Theorem entail:

$$\Gamma, \Pi \vdash A \rightarrow B \tag{37}$$

(34) and Monotonicity entail:

$$\Gamma, \Pi \vdash A \tag{38}$$

(37), (38) and MP then entail:

$$\Gamma, \Pi \vdash B \tag{39}$$

<u>CLASSICAL CONSEQUENCE:</u> If A follows from Γ by classical logic, then $\Gamma \vdash A$

Let us suppose that:

$$A \text{ follows from } \Gamma \text{ by classical logic.} \tag{40}$$

(40) and PC entail:

$$\varnothing \vdash \Gamma \rightarrow A \tag{41}$$

(41) and the Monotonicity entail:

$$\Gamma \vdash \Gamma \rightarrow A \tag{42}$$

PC entails:

$$\Gamma \vdash \Gamma \tag{43}$$

(42), (43) and MP collectively entail:

$$\Gamma \vdash A \tag{44}$$

<u>MODUS PONENS:</u> $\Gamma, A \Rightarrow B, A \vdash B$ where $\Rightarrow$ is either the material or counterfactual conditional.

This proof precedes in two steps—one for the material conditional and the other for the counterfactual conditional. Let us begin with the material conditional. Classical Consequence, having been just established, entails:

$$\Gamma, A \rightarrow B, A \vdash A \tag{45}$$

In addition, Classical Consequence entails:

$$\Gamma, A \rightarrow B, A \vdash A \rightarrow B \tag{46}$$

(45), (46) and MP entail:

$$\Gamma, A \rightarrow B, A \vdash B \tag{47}$$

This establishes the material version of Modus Ponens. The counterfactual version requires additional steps. First, $MP\Box$ entails:[20]

$$\varnothing \vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \rightarrow B) \tag{48}$$

Due to Monotonicity, this entails:

$$\Gamma, A \mathrel{\Box\!\!\rightarrow} B, A \vdash (A \mathrel{\Box\!\!\rightarrow} B) \rightarrow (A \rightarrow B) \tag{49}$$

Classical Consequence, in turn, entails:

$$\Gamma, A \mathrel{\Box\!\!\rightarrow} B, A \vdash A \mathrel{\Box\!\!\rightarrow} B \tag{50}$$

(49), (50) and MP collectively entail:

$$\Gamma, A \mathrel{\Box\!\!\rightarrow} B, A \vdash A \rightarrow B \tag{51}$$

Classical Consequence entails:

$$\Gamma, A \mathrel{\Box\!\!\rightarrow} B, A \vdash A \tag{52}$$

And, finally, (51), (52) and MP entail:

$$\Gamma, A \mathrel{\Box\!\!\rightarrow} B, A \vdash B \tag{53}$$

COUNTERFACTUAL DEDUCTION: If $\Gamma, A \vdash B$ then $\Gamma \vdash A \mathrel{\Box\!\!\rightarrow} B$

[20] Note that this is the only instance where MP$\Box$ is employed within this proof; it is not needed to prove any other axiom that Hawthorne and Yli-Vakkuri rely upon.

Let us suppose:

$$\Gamma, A \vdash B \tag{54}$$

(54) and Deduction Theorem entail:

$$\Gamma \vdash A \rightarrow B \tag{55}$$

(55) and Closure entail:

$$\Gamma \vdash (A \mathrel{\Box\!\!\rightarrow} A) \rightarrow (A \mathrel{\Box\!\!\rightarrow} B) \tag{56}$$

An application of Reflexivity is:

$$\varnothing \vdash A \mathrel{\Box\!\!\rightarrow} A \tag{57}$$

Due to Monotonicity, this entails:

$$\Gamma \vdash A \mathrel{\Box\!\!\rightarrow} A \tag{58}$$

(56), (58) and MP then entail:

$$\Gamma \vdash A \mathrel{\Box\!\!\rightarrow} B \tag{59}$$

DEDUCTION THEOREM

As before, Deduction Theorem is an axiom in both systems under consideration and follows trivially from itself.

Therefore, all Yli-Vakkuri and Howthorne's axioms follow from Williamson's when appended to Deduction Theorem. Because each set of axioms can be derived from the other, the two systems are equivalent. As I mentioned at the outset, Williamson independently established that his system is equivalent to a **T** modal logic. It follows that Yli-Vakkuri and Howthorne's axioms are equivalent to appending **T** to Deduction Theorem. Everything provable in their system is provable in the conjunction of Deduction Theorem with **T**; nothing which cannot be proven in their system can be proven in the conjunction of Deduction Theorem with **T**.