Incompleteness Theory of Euclidean Possible-Worlds Semantics and Resolution of the Surprise Quiz Paradox

---

I
<u>Abstract</u>

Our best attempt at formalizing modal intuitions in a logical system fails. Despite previous work indicating that we have formalized modal intuitions into a complete and consistent logical system, we have not. Rather, any system of modal logic formalized within a Euclidean possible-worlds semantics (such as through S5) is either incomplete—that is, it fails to prove a true modal proposition—or it is inconsistent—that is, it proves inconsistent modal propositions.

I demonstrate this incompleteness in modal logic by engrafting a correspondent metalanguage—first an epistemological and then a formal one—onto a Euclidean possible-worlds semantics. In each of these metalanguages—which are governed by restrictive rules for formulation and derivation to ensure truth-functional correspondence to modal-logic systems formalized within the underlying semantics—I derive an undecidable and then a contradictory modal proposition. Specifically, I employ the corresponding formulation and derivation of *mere possibilia*—i.e., contingently non-actual propositional objects. In corresponding the metalanguages to the underlying semantics, I thus demonstrate the incompleteness of any system of modal logic formalized within Euclidean possible-worlds semantics.

In taking account of this poverty of modal logic, the unintuitive conclusion of the Surprise Quiz Paradox no longer follows from its premises. The conclusion of the paradox that "it is impossible for there to be a sur-

prise quiz" is thus merely erroneous. Critically, as accounted-for in this resolution of the Surprise Quiz Paradox, the poverty of modal logic extends *only* to this erroneous conclusion, preserving modal intuitions. I demonstrate this resolution first in the epistemological and then in the formal metalanguage.

I conclude with prescriptions for the way forward, and I propose that no more than three possible worlds may be semantically tenable.[1]

II
Background

Modal intuitions are intuitions about how to work with statements of possibility and necessity. Statements of *a posteriori* possibility include "it is possible that on Friday night, I go bowling instead of seeing a movie." Statements of *a posteriori* necessity (we think) include "there necessarily was a big bang."

Possible worlds are a way of making sense of our *a posteriori* modal intuitions. For example, a possible-worlds framework underpins important formal logical systems expressing our modal intuitions. These systems enable philosophers to construct *a posteriori* arguments that rely on modal intuitions and to say that the arguments proceed by logical force.

For our purpose here, a "possible world" is a logically possible world: just a collection of laws and circumstance that actually could be—viz., that is possible. This separates it from an impossible world, in which a collection of laws and circumstance simply cannot be—like a world in

[1] Special thanks is due to Eric Dietrich, my philosophy professor. Without his unique tutelage and insight, this paper would not be. May philosophy departments overflow evermore with such rigorous philosophers who seek the good as truth and beauty.

which the law of gravity is true but a random half of things fall down and the other half float up. We say that "it is possible that on Friday night, I go bowling instead of seeing a movie" because there are worlds—collections of laws and circumstance—in which I go bowling instead of seeing a movie on a particular Friday night. We say that "there necessarily was a big bang" because (we think) there is no world—no collection of laws and circumstance—in which there never was a big bang.

The actual world is whatever world is actually in existence right now. An essential property of an actual world, we intuit, is that it is unique—there can be only one actual world at a time.

This essay concerns a formal possible-worlds framework that has the property of Euclidean accessibility relation, which forms the interpretational syntax of S5, the most robust formal system of modal logic. "The property of Euclidean accessibility relation" is a technical designation. Metaphorically, it means that a set of possible worlds a philosopher decides to talk about in the same argument are all on the "same metaphysical page" with respect to what kinds of particular things can be possible or necessary in each of the possible worlds. More particularly, it means that if possible world *v* is accessible from possible world *w*, every possible world accessible from *w* is also accessible from *v*, and vice-versa. Put another way, the property of Euclidean accessibility relation comprises three more primitive properties: reflexivity—i.e., *w* is accessible from *w*—symmetry—i.e., where *v* is accessible from *w*, *w* is accessible from *v*—and transitivity—i.e., where *v* is accessible from *w* and *u* from *v*, *u* is accessible from *w*.

To say that one possible world is "accessible" from another and vice-versa means that the same *modal* propositions in both take the same truth values. For instance, if two possible worlds are accessible from each other, it cannot be true that in one "it is impossible to fly faster than the speed of light" while the same statement is false in the other. A possible-worlds framework with Euclidian accessibility relation is a complete graph of possible worlds. In such a graph, every possible world is represented by a vertice, and all vertices are connected with one another. Each connection—or edge—represents an accessibility relation. With any more than three possible worlds, the number of accessibility relations exceeds the number of worlds. At the metaphorical center of this graph (to the extent that a complete graph can have a center) is our world—the actual world—in which we exist as we sit down at our desks.

III
Incompleteness Through Epistemological Metalanguage

A. The Mind-Experiment Test of the Logical Possibility of a World

Logical possibility is the broadest alethic modality. Simply, a proposition of logical possibility is true if and only if it can be asserted without implying a logical contradiction. Thus, to know that a world is logically possible is to know that a collection of laws and circumstance is logically possible.

We now set forth a corresponding metalinguistic model of logical possibility within restrictive epistemological rules: to know that a world is logically possible is to know that you are able, without contradiction, to do a mind experiment. Even if *you don't actually do it*, you must know that you are able to do it without contradiction, or else you cannot know that a collection of laws and circumstance really is possible and therefore constitutes a possible world. If you know that you are able, without contradiction, to do the mind experiment, you then know that the collection of laws and circumstance constitutes a possible world.

The mind experiment is this: first, suppose that you have limitless capacity to apprehend the implications of all laws and circumstance and to detect when some permutation of laws or circumstance would result in contradiction; next, in your mind, imagine moving from the actual world to a possible world and making that possible world the actual world. What do we mean by "move to" the possible world? We don't mean, in our minds, putting ourselves in that possible world—we only do that if that possible world is so defined as to include us. Simply, we mean that we suppose the possible world actually exists.

B. The First World Blanks out of Possibility

But if you do that—in your mind experiment, move from the actual world (call it the "first world") to a possible world, and then you make that possible world the new actual world (call it the "second world") then, in your mind experiment, what is the first world? In short, when you suppose that the second world is the actual world, what do you then suppose the first world is?

The natural answer seems intuitive—but the natural answer is flawed.

The natural answer is that you just switch the worlds. That is, in the mind experiment, the first world becomes just another possible world. Whereas before, the first world was actual and the possible world was possible but not actual, you just switch. The possible world becomes the actual world and the first world becomes the world that is possible but not actual. Just switch. Right? Wrong.

Consider this: in the mind experiment, after we move to the second world, what can we be sure the first world is *not*?

In the mind experiment, we can be sure the first world is not the actual world because there can be only one actual world—viz., an essential property of an actual world, we intuit, is that there is only one at a time. When we move, the second world becomes the actual world. So in the mind experiment, the first world surely is not the actual world. Simple enough.

This essay will argue for the following additional conclusion: for all we know, if in the mind experiment the first world is not the actual world, then we have the strongest justification to believe that in the mind experiment, the first world—the world we are in when we do the mind experiment—is an impossible world.

Why? Is that not too strong? Did we not need just to say that the first world is no longer the actual world? Is it not too strong to say that the first world is not even possible anymore?

First, we should clarify the conclusion: we do not say that the first world really is impossible—obviously the first world is still the actual world. After all, we're in it, doing the mind experiment.

But, the mind experiment is implicitly premised on the first world being an impossible world. As evidence, we can run the mind experiment an infinity of times, and each time we do, the first world we're in as we do the mind experiment is never the actual world in the mind experiment. We can even change everything about the second world and wait an infinity of time while everything about the first world changes and then run the mind experiment an infinity of times again; no matter what we do or what happens, in the mind experiment, the world we are in when we do the mind experiment—the first world—is not the actual world. This is inductive evidence constituting the strongest justification for our believing that, for all we know, in the mind experiment, the first world is an impossible world. If in the mind experiment the first world is not the actual world an infinity of times and in an infinity of conditions, calling the first world still "possible" in our mind experiment is disingenuous. We'd just be using a label without a meaning.

The mind experiment's implicit premise that the first world is impossible stems from the essential feature of the mind experiment: it must test whether a world that is not the actual world is a possible world. Given that, if the mind experiment is designed properly to provide us the knowledge we seek from it, we must always assume that the first world—the world we're in when we do the mind experiment—is not the actual world. That is simply because, in the mind experiment properly designed, we must assume that the second world—the one we want to test—is the actual world, and an essential property of an actual world, we intuit, is that there is only one at a time.

But is it really true that we are just "labeling" the first world as possible when we do the mind experiment and suppose that the first world is not the actual world for but a moment? In the mind experiment itself, would not the test of whether the first world remains possible be the same? A mind experiment? If so, what stops some possible conscious being in the second world—called the "second man"—from describing and naming the first world and saying "that world is possible" by doing his own mind experiment—called the "reverse mind experiment"—moving from his world to ours, imagining our world as an actual world? Nothing. Unless rejecting the possibility of other conscious beings, the possibility of this second man's mind experiment is assured—indeed, it is a property of the second world: for our purpose, this means that the second mind-experimenting man must "exist" *either* in the second world or another possible world accessible both from the second world and the first.[2]

On the basis of the second man's mind experiment, the second man would have justified belief that the first world is possible. In moving to the second world in our own mind experiment, we must suppose that a second man does his own mind experiment to move to the first world.

On a separate basis, we would know that the justified belief that the second man would develop through his mind experiment—that the first world is possible—is a true belief. Our separate basis is that we are in the first world as we do our mind experiment, so we know that the first world really is the actual world; therefore, we know that it is possible.

In our mind experiment, the second man still could not *know* that the first world is possible even though it is, he would believe it, and he would have justification for believing it. That is because of the Gettier Problem.

The Gettier Problem arises when there is no connection between the truth of a proposition one believes and his justification for believing it. In sum, the Gettier Problem is about having the wrong good reason for a true belief. Edmund Gettier provides this critical example:

> Let us suppose that Smith has strong evidence for the following proposition:
>
> (f) Jones owns a Ford.

[2] Remember, it is not that any of these mind experiments are ever done. It is that they can be done without contradiction. Doing them is the test of whether they can be done without contradiction.

> Smith's evidence might be that Jones has at all times in the past within Smith's memory owned a car, and always a Ford, and that Jones has just offered Smith a ride while driving a Ford. Let us imagine, now, that Smith has another friend, Brown, of whose whereabouts he is totally ignorant. Smith selects three place-names quite at random, and constructs the following three propositions:
>
> (g) Either Jones owns a Ford, or Brown is in Boston;
>
> (h) Either Jones owns a Ford, or Brown is in Barcelona;
>
> (i) Either Jones owns a Ford, or Brown is in Brest-Litovsk.
>
> Each of these propositions is entailed by (f)[through the rule of disjunctive addition[3]].
>
> Imagine that Smith realizes the entailment of each of these propositions he has constructed by (f), and proceeds to accept (g), (h), and (i) on the basis of (f). Smith has correctly inferred (g), (h), and (i) from a proposition for which he has strong evidence. Smith is therefore completely justified in believing each of these three propositions. Smith, of course, has no idea where Brown is.
>
> But imagine now that two further conditions hold. First, Jones does not own a Ford, but is at present driving a rented car. And secondly, by the sheerest coincidence, and entirely unknown to Smith, the place mentioned in proposition (h) happens really to be the place where Brown is. If these two conditions hold then Smith does not know that (h) is true, even though (i) (h) is true, (ii) Smith does believe that (h) is true, and (iii) Smith is justified in believing that (h) is true.

Edmund Gettier, 23 Analysis 121, 122–23 (1963).[4]

In the mind experiment, the second man's own mind experiment would serve as the basis for the second man's justified true belief that the

[3] The rule of disjunctive addition instructs, for example, that if it is true that "your name is Saul," then it is true that "your name is Saul" or that "the moon is made of blue cheese"—the second disjunct can be anything, regardless of whether it is true. That is because only one disjunct in a disjunction needs to be true to make the whole disjunction true.

[4] Let us coyly add that later, Smith learns that Jones owns no Ford. When Smith learns this, he is surprised.

first world is possible.

In the mind experiment, the second man's justified true belief that the first world is possible, however, is not knowledge, because of the Gettier Problem. His justified true belief suffers the Gettier Problem because he has the wrong justification for his true belief that the first world is possible. He has the wrong justification for that true belief because his justification is inconsistent with ours. Our basis is that we are in the first world as we do our mind experiment, so we know that the first world really is the actual world—therefore, we know that it is possible. In the mind experiment, however, that basis is one that the second mind-experimenting man must reject because in the mind experiment, the second world and necessarily not the first world is the actual world.[5]

We can demonstrate the second man's Gettier Problem by extending Gettier's example quoted above. The example below tracks Gettier's own language:

Suppose in our mind experiment that Smith is the second mind-experimenting man. Suppose further that Smith has strong evidence for the following proposition:

(a) The first world is possible but not actual.

Smith's evidence is that he has conducted his own mind experiment from the second world—the actual world in his mind experiment—to the first world. Let us imagine, now, that Smith, quite at random, selects two collections of laws and circumstance from a giant hat without looking,

---

[5] Formally, if we were to represent our basis, it would derive from Axiom (B) of S5 ($A \to \Box\Diamond A$; viz., "what is actually true is necessarily possible."). If we were formally to represent the second man's basis, it would derive from Axiom (5) of S5 ($\Diamond A \to \Box\Diamond A$; viz., "what is possible is necessarily possible.").

puts each of those collections into their own bins—again, without looking—and then names each bin "world *g*" and "world *q*." Within a mind experiment, he then constructs the following two propositions, (b) and (c):

(b) Either "the first world is possible but not actual" or "world *g* is the actual world."

(c) Either "the first world is possible but not actual" or "world *q* is the actual world."

Each of these propositions is entailed for Smith by (a) through the rule of disjunctive addition. Imagine that Smith realizes the entailment of each of these propositions he has constructed by (a) and then proceeds to accept (b) and (c) on the basis of (a). Smith has correctly inferred (b) and (c) from a proposition for which he has strong evidence. Smith is therefore completely justified in believing each of these two propositions. Smith, of course, has no idea what laws and circumstance he picked from the giant hat.

But imagine now that two further conditions hold. First, Smith himself "exists" as part of a mind experiment and, unknown to him, the first world is really the actual world, not the second world he believes is the actual world. And second, by the sheerest coincidence, and entirely unknown to Smith, the collection of laws and circumstance he picked from a giant hat and named world *q*, which he said was the actual world in proposition (c), matches in all ways the actual world; therefore, by the rule of disjunctive addition, proposition (c) is true.

If these conditions hold, then Smith does not know that (c) is true, even though (i) (c) *is* true, (ii) Smith does believe that (c) is true, and (iii) Smith is justified in believing that (c) is true.

Why does it matter that the second man's justified true belief that the first world is possible does not rise to knowledge? Isn't his justified true belief strong enough? No.

The second man's Gettier-flawed justified true belief which falls short of knowledge is not enough for us to conduct without contradiction our mind experiment and know that on the basis of that mind experiment, the first world is still possible. If we want to know that the first world is still possible on the basis of the mind experiment, unless we want to suffer the Gettier Problem too, the second man really must have the right justification and know, in our mind experiment, that the first world is possible. Otherwise, if, on the basis of the mind experiment, we try to accept the second man's justified true belief that the first world is possible as the basis of our knowledge of the same proposition, we will have the wrong justification, too—and our belief, whether justified or true, will suffer the Gettier Problem.

In the mind experiment, when we move from the first world to the second world, the first world blanks out of possibility. For a world to "blank out of possibility" means that, on the basis of the mind experiment, we have the strongest justification to believe that it is an impossible world.

C. Blanking out of Possibility All the Possible Worlds—Except for One Last, Odd Straggler

Of course, this does not mean that we know that the actual world is impossible. The actual world and the first world are not the same thing; the latter is a possible world corresponding to the actual world. In this vein, nothing stops us from doing a mind experiment to another second world, one also corresponding to the actual world, to show that the actual world is still possible. Ultimately, this won't succeed, however, because within any single mind experiment, we can do an infinite succession of mind experiments, and from this, as explained below, it would follow that we have the strongest justification to believe that either the actual world is necessary or cannot be.

In the mind experiment, we can move from the first world to the second world, blanking out of possibility the first world; we can then have the second man move to a third world, blanking the second world out of possibility. And on and on it goes.

The third world becomes "second world prime" and the second world becomes "first world prime." The second world prime may be any world. First world prime corresponds to the original second world. The second man—who is the same second man—is still called the "second man." The second man's move to second world prime is "mind experiment prime."

"Second man prime" will be he who is, to the second man in mind experiment prime, the second mind-experimenting man.

The "reverse mind experiment prime" is the mind experiment of the second man prime to first world prime within the second man's mind experiment prime.

The possibility of a second man, without contradiction, doing mind experiment prime to move to second world prime is a property of first world prime. And so, in moving to the second world (which corresponds to first world prime) in our own mind experiment, we must suppose that the second man does his own mind experiment to move to second world prime.

What do we mean by "move to" second world prime from first world prime? We don't mean, in our minds, putting ourselves in second world prime—we only do that if second world prime is so defined as to include us. We don't mean, in the second man's mind, putting him in second world prime. We only do that if second world prime is so defined as to include the second man. We mean to suppose in our mind experiment that first world prime actually exists and then to suppose that the second man does mind experiment prime to second world prime. The second man thereby supposes second world prime actually exists.

For the second man in mind experiment prime, the mind experiment of second man prime to first world prime—reverse mind experiment prime—will suffer the Gettier Problem. The second man in mind experiment prime will therefore have the strongest justification for believing that first world prime is impossible.

But then, in mind experiment prime, the second man must suppose that second man prime begins mind experiment double prime: and on and on we go, until all worlds are blanked out of possibility but one. (All this, by the way, is still going on within *our* mind experiment.)

After we have in our mind experiment blanked out of possibility all the worlds comprising every possible permutation of laws and circumstance and there are no others to move to, the last world we wind up on gets to keep its possibility. That is because we will have no other second world prime to move to. So, the last world we wind up on will be the last possible world left in our mind experiment.

Being the last possible world left in our mind experiment, we will have the strongest justification for believing that all the laws and circumstance that constitute it become necessary—that is, true in all possible worlds—such that the last possible world becomes necessary. Either this last, necessary world corresponds to the actual world, in which case the actual world is necessary—which is to fail to prove any number of true propositions of *mere possibilia*—or this last, necessary world is distinct from the actual world, implying that more than one world must be actual—which is to prove contradiction.

IV
Incompleteness Through Formal Metalanguage

Here, we adapt Cantor's diagonal argument to a correspondent formal metalanguage.

Suppose that for every proposition A in S5-based logic L at time T^{α}, there is a corresponding metalinguistic form A^{α} {… [Y/N], [Y/N], [Y/N], …}. For any A^{α}, "Y", "N", or "[Y/N]" is assigned for each world w based on whether proposition A takes an object that exists in w. (For example, where A takes an object that exists in no possible worlds, A^{α} {… N, N, N, …}.)

Suppose further that only the rules of reflexivity and symmetry govern this form, such that A^1 {… x, x', …} if and only if A^2 {… y, y', …}, but only when world x = world y, world x' = world y', and so forth.

We should then suppose that given any A, its corresponding form A^{α} and any form derived by the above rules should take the same truth value that A takes in L.

Now, let us construct a model of *mere possibilia* at time T^{α}. A takes an object that is possible but false in the actual world. Its corresponding form—called the primary form—is therefore A^{α} {… a, b, c, d, …} such that for at least one world w, N, and for at least one world w', Y. Further, for every such corresponding form of *mere possibilia* for worlds at time T^{α}, there is a complimentary form for the same worlds at time T^{ω} at which the propositional object of A—and only that object—ceases to be in one world but remains possible in the others. Only through the primary form and its complimentary form do we fully express the contingency of *mere possibilia*: at one time something may be but not necessarily and accordingly may be at one time but not at a future time; and stated precisely and completely, the propositional object of *mere possibilia* may be true in one world and false in another and may be true in one world and false at a future time in that same world.

We now represent this model so that it may express every permutation of *mere possibilia* in unbounded sets of possible worlds and propositional objects. (We should expect such project to be successful if L is complete and consistent.) Thus, we construct a graph with vertical and horizontal axes extending infinitely in both directions, a central axis to the right of which are the primary forms for worlds at time T^{α} and to the left of which are the complimentary forms for worlds at time T^{ω}, coordinates defining vertices at which worlds intersect with a propositional object, and sets of such coordinates designated with sequential assignments of the indexical A (with each indexical A assigned to a proposition *A* in L). Set each indexed primary form mutually to imply a corresponding form indexed equidistantly from the central axis. Thus, A^{n} mutually implies A^{-n}, A^{n+1} mutually implies $A^{-(n+1)}$, and so forth.

A proposition of *mere possibilia* is agnostic as to which worlds contain the propositional object and which do not, so long as at least one does and at least one does not. Accordingly, to represent every permutation of *mere possibilia* within the graph, the propositional object exists and fails to exist in each world at least once. To represent this, a diagonal within the graph is constructed like so:

	…	…	
{ … N, Y, [Y/N], [Y/N], … }	$n+2$	$A^{-(n+2)}$	
{ … [Y/N], N, Y, [Y/N], … }	$n+1$	$A^{-(n+1)}$	
{ … [Y/N], [Y/N], N, Y, … }	n	A^{-n}	
	…	…	
… $n+3$ $n+2$ $n+1$ n …	T^{ω}	T^{α}	… n $n+1$ $n+2$ $n+3$ …
	…	…	
	A^{n}	n	{ … N, Y, [Y/N], [Y/N], … }
	A^{n+1}	$n+1$	{ … [Y/N], N, Y, [Y/N], … }
	A^{n+2}	$n+2$	{ … [Y/N], [Y/N], N, Y, … }
	…	…	

The diagonal is highlighted in gray. According to the agnosticism of *mere possibilia*, the propositional object may either exist or fail to exist in worlds outside of the diagonal, as represented by the value "[Y/N]".

Now, derive an instance of *mere possibilia* expressible in the corresponding form but undecidable in L: A^{ω} {… *aa*, *bb*, *cc*, *dd*, …} such that *aa* is Y where *a* is N and N otherwise, *bb* is Y where *b* is N and N otherwise, and so on. We should expect that this function will yield a legal instance of *mere possibilia* because in at least one world the propositional object will exist and in at least one other the propositional object will fail to exist.

However, this instance of *mere possibilia* cannot be expressed in our graph that should contain every permutation of *mere possibilia*. A^{ω} cannot be A^{n} because the forms will be inconsistent at vertices n, n and n, $n + 1$. A^{ω} cannot be A^{n+1} because the forms will be inconsistent at vertice $n + 1$, $n + 2$. A^{ω} cannot be A^{n+2} because the forms will be inconsistent at vertice $n + 2$, $n + 3$. And so on ad infinitum. The same infinite regress results no matter which indexical A is considered first. Going backwards works no better, as any rearward form will imply a complimentary form that will imply a primary form eliminated by the regress. So, for example, A^ω cannot be A^{-n} because A^{-n} implies A^n and by the previous demonstration, A^ω cannot be A^n; and so on and so forth, also ad infinitum.

Either this instance cannot be proven in L—in which case, L is incomplete—or the instance is impossible in L—in which case, L yields contradiction.

V
Resolution of the Surprise Quiz Paradox

The Surprise Quiz Paradox:

> A teacher announces that there will be a surprise quiz next week. A student objects that this is impossible: "The class meets on Monday, Wednesday, and Friday. If the quiz is given on Friday, then on Thursday I would be able to predict that the quiz is on Friday. It would not be a surprise. Can the quiz be given on Wednesday? No, because on Tuesday I would know that the quiz will not be on Friday (thanks to the previous reasoning) and know that the quiz was not on Monday (thanks to memory). Therefore, on Tuesday I could foresee that the quiz will be on Wednesday. A quiz on Wednesday would not be a surprise. Could the surprise quiz be on Monday? On Sunday, the previous two eliminations would be available to me. Consequently, I would know that the quiz must be on Monday. So a Monday quiz would also fail to be a surprise. Therefore, it is impossible for there to be a surprise quiz."

Roy Sorensen, *Epistemic Paradoxes*, in STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2017).

A. Epistemological Resolution

Broadly, because of the Gettier Problem in reverse mind experiments, the student cannot suppose *in his mind experiment*, which he conducts in the actual world prior to Monday, that he will know on Thursday that the quiz was not on Monday. In his mind experiment, for all he knows, then, on Thursday, the quiz already happened and it makes no difference that the quiz cannot be on Friday. (It would make no difference because the quiz only happens once.) Crucially, this leaves undisturbed our intuition that the surprise quiz cannot be on Friday: it just makes that intuition immaterial to whether the quiz can be on Monday or Wednesday, for all the student knows before Monday.

The student formulates his argument that a surprise quiz is impossible in the actual world. The actual world is one prior to Monday—let's say Sunday—in which the student has learned from his teacher that there will be a surprise quiz on the subsequent Monday, Wednesday, or Friday.

The student's argument is about what is possible and what is necessary; the student must therefore conduct mind experiments. Specifically, the student's argument proceeds as follows:

First, the student moves to a second world.

The second world is one in which Tuesday has arrived and there was no quiz on Monday: "[O]n Tuesday I would know . . . that the quiz was not on Monday (thanks to memory)." The student conducts this move without contradiction.

Next, the student would like to conclude that the quiz cannot be on Wednesday because it didn't happen on Monday, but before he does that, he has to demonstrate that *on Tuesday* the quiz cannot be on Friday. Put differently, the student must demonstrate that if the quiz didn't happen on Monday, on Tuesday it is the case that the quiz cannot be on Friday and so would have to be on Wednesday and thus not a surprise. So:

Second, the student moves from the second world of Tuesday to a second world prime of Thursday.

Second world prime is one in which Thursday has arrived and the quiz has yet to happen: "[O]n Tuesday [of the second world] I would know that the quiz will not be on Friday (thanks to the previous reasoning)," the previous reasoning being, "If the quiz is given on Friday [that is, if it has yet to be given on Thursday], then on Thursday [of second world prime] I would be able to predict that the quiz is on Friday." The student thus moves from the second world to second world prime. The student conducts this move without contradiction.

But, the student may move back from second world prime of Thursday to the second world of Tuesday via reverse mind experiment prime. (In reverse mind experiment prime, the second world is also called "first world prime.") This reverse mind experiment will fail to provide knowledge that from second world prime of Thursday, the second world of Tuesday (that is, first world prime) is still possible. So, for all second man prime of reverse mind experiment prime knows, first world prime is impossible. For all second man prime knows, therefore, a world in which Tuesday has arrived and there was no quiz on Monday is impossible.

This all happens, of course, in the student's own mind experiment. Because the student on the basis of his own mind experiment cannot know *both* that it is possible that a surprise quiz does not happen on Monday and cannot happen on Friday, he cannot know that a surprise quiz is impossible on Wednesday. Following this, because the student cannot know *both* that a surprise quiz cannot happen on Wednesday and cannot happen on Friday, he cannot know that a surprise quiz is impossible on Monday. Crucially, he still can know that a surprise quiz is impossible on Friday—but only on Friday—because anyone from any world cannot move without contradiction to a world in which Friday has arrived, the quiz to be either on that Friday or a prior day has not yet happened, but the quiz is still a surprise. Such a world is necessarily impossible.

Does this resolution really work? The Gettier Problem, the student might object, involves justified true belief. Being true, is not justified true belief good enough for the student to hold out his conclusion as true, even if he does not know it? Apparent tension notwithstanding, the answer is no. The apparent tension is that typically, the Gettier Problem involves justified belief in something that is true; but here, the student's conclusion as to Monday and Wednesday is false. But this apparent tension can be explained.

Here, the student has justified true beliefs that certain worlds are possible together. These beliefs are intermediary to the student's ultimate conclusion. The truth of the student's intermediary justified beliefs fails to transfer to the student's ultimate conclusion based off of those beliefs, even if the justification does. The student's ultimate conclusion may at best, then, represent justified but false belief. By contrast, knowledge that the same certain worlds are possible together would transfer truth, by implication, to an ultimate conclusion based off of that knowledge. Why exactly this is may be a topic of further inquiry. It likely involves how the right justification for an antecedent may be key to the overall truth of a conditional if-then statement when the statement's antecedent is true. The Surprise Quiz Paradox may represent an especially strong Gettier Problem, showing just how big a difference full-fledged knowledge, complete with the right justification, can make.

B. Formal Resolution

Consider the resolution without epistemological reference to an individual knower. Proffer the propositional object "a quiz." Let A^n represent the proposition "it is possible that the student remembers that the quiz was not on Monday" and A^{n+1} represent the proposition "it is possible that the student remembers that the quiz was not on Wednesday." Setting the world at n, n to Monday and the world at n, $n+1$ to Wednesday, the corresponding metalinguistic form to the student's essential suppositions are: A^{n} { N, [Y/N] } and A^{n+1} { [Y/N], N }. A^{n+1} is undecidable in L as it is an instance of the undecidiable A^{ω} defined above. Put differently, applying the function for A^{ω} to A^{n} yields a form consistent with A^{n+1}. Because the student's second essential supposition is undecidiable in L, the student's argument is invalid.[6]

VI
The Way Forward

We should decide whether our results here tell us either (1) that we should discover a new way of structuring our analysis of possibility and necessity without possible worlds or (2) that we should accept new philosophic conclusions flowing from the incompleteness of our possible-worlds based analysis of possibility and necessity. Or both. It would appear at first that the resolution of the Surprise Quiz Paradox is a splendid philosophic conclusion revealed from our otherwise savage demolition of the revered philosophic edifice of possible worlds. I would doubt, however, that the resolution of this particular longstanding paradox exhausts the philosophic conclusions we can mine from the incompleteness of our most developed analysis of possibility and necessity. I suspect, for example, that we may solve all of Zeno's paradoxes.

[6] In achieving this solution, we instantiate the metalinguistic model of *mere possibilia* in a certain way: we insert a single, spatiotemporally agnostic propositional object into a graph at all vertices and then spatiotemporally sequence the vertices. In a similar way, likely all of Zeno's paradoxes may be solved.

Alternatively, perhaps a new edifice can replace possible-worlds semantics. I would doubt, however, that we know precisely what such a new edifice—that is, a new formal system for making sense of modal propositions—will get us that would be much better than what we gain by tearing our existing one down. For example, will the Surprise Quiz Paradox reemerge? Will we lose new philosophic discoveries that formal modal systems have obscured from us?

Most important for us to consider, I would think, in deciding between our two alternatives is whether possibility and necessity are actually distinct phenomena. Perhaps in our haste to suppose that the field of philosophy makes any progress at all, we have forgotten that deep philosophic questions such as divine or material determinism remain. For my part, I suspect that there is such a thing as possibility, but that it is a mystery.

We face not the necessary prospect that there is *no* possibility and necessity. The incompleteness of modal logic emerges after an infinite regress. We should first probe the structure of infinity and the regress to determine whether some notion of possibility and necessity limited in some feature resolves the incompleteness. In so doing, we should manipulate the contours of possibility and necessity to discover whether these categories mask distinctions that are both finer and more important. Perhaps our study of quantum mechanics will be illuminating in this regard. Or, perhaps not. Time will tell. (Or, perhaps it will not.)

Perhaps we should explore whether our results here to do more than inform the structure of our formal means for representing possibility and necessity and instead tell us something particular about what *is* possible and necessary. For example, if the number of accessibility relations and possible worlds were identical in a trinity of three each, would the regress still obtain? I think not. Would the regress still obtain if David Lewis is right and every possible world really is an actual world—somewhere?

Perhaps we should also explore whether an important distinction exists between something possibly being and something possibly not being. We seem, in building our possible-worlds framework—and, in fairness, simply reflecting upon our modal intuition—to see no such distinction. This seems so, I would imagine, because we do not live in possible but non-actual realities and from here, any non-actual reality looks and feels the same—and so, we had supposed, any such reality should be formalized the same way in modal logic: as possible but neither actual nor necessary.

Well, if possibility and necessity are real phenomena, then we should have understood our blunder, obvious in hindsight (as many things in such sight of course are): we failed to consider the nature of non-actuality. A distinction between something that possibly is and possibly is not must consider the nature of non-actuality, and with any luck—and perhaps with help from quantum mechanics (or perhaps not)—we will consider such nature in the right way. With aid from our already existing and reasonably well-developed intuitions of conditional probability, we can perhaps start speculating as to the nature of non-actuality. In any event, it is certain that bold philosophic action—to include speculations and conjectures of all sorts—will be necessary to build our results today into philosophic (or even scientific) progress tomorrow.

And in so acting, we must cease the pleasant tendency to see progress in ever-more-dense markings and formalisms that seem to have captured philosophic prestige from the Aristotle-like dreamers who must now come to our rescue.

Sean Ross Callaghan June 19, 2017

sean.ross.callaghan@gmail.com