



## Philosophical Review

---

The Logic of What Might Have Been

Author(s): Nathan Salmon

Source: *The Philosophical Review*, Vol. 98, No. 1 (Jan., 1989), pp. 3-34

Published by: Duke University Press on behalf of Philosophical Review

Stable URL: <https://www.jstor.org/stable/2185369>

Accessed: 24-03-2020 13:02 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*Duke University Press, Philosophical Review* are collaborating with JSTOR to digitize, preserve and extend access to *The Philosophical Review*

## THE LOGIC OF WHAT MIGHT HAVE BEEN\*

Nathan Salmon

In earlier work I argued (following Hugh Chandler) that the conventionally accepted system *S5* of (first-order) modal propositional logic, and even the weaker system *S4*, embody an invalid pattern of modal reasoning; they are fallacious systems for reasoning about what might have been.<sup>1</sup> I argued, in fact, that the characteristic *S4* axiom schema, " $\Box\phi \supset \Box\Box\phi$ "—or equivalently, the principle that for any necessarily true proposition *p*, the proposition that *p* is necessarily true is itself necessarily true—is not only not logically true, some instances are in fact untrue. I argued, that is, that for some necessary truths *p*—for example, that a certain table does not originate from a certain hunk of wood—the fact that *p* is necessary cannot itself be correctly deemed necessary. Instead, although any such proposition *p* is necessary, the claim that *p* is necessarily necessary is untrue, and indeed some claim of the form " $\Box\Box. . . \Box p$ " is altogether false.

While some of my audience have found these arguments against *S4* modal logic persuasive, many have found them unconvincing. I have repeatedly encountered two particular objections, which are probably best regarded as two parts of a single objection. This objection, however, betrays a serious misunderstanding of my position, or a failure to appreciate the full force of my (Chandleresque) arguments, or both, and is based on a confusion among concepts central to the foundations of contemporary semantics for

---

\*This paper was presented to an international conference on Meaning and Natural Kinds at the Inter-University Centre of Postgraduate Studies in Dubrovnik, Yugoslavia, September 1986. It has benefitted from a discussion with Timothy Williamson, and from comments by Hugh Chandler, Graeme Forbes, and the anonymous referees.

<sup>1</sup>Hugh Chandler, "Plantinga and the Contingently Possible," *Analysis* 36 (1976), pp. 106–109. For my renderings of Chandleresque arguments, see *Reference and Essence* (Princeton, N.J.: Princeton University Press and Oxford, England: Basil Blackwell, 1981), section 28, pp. 229–252; "Impossible Worlds," *Analysis* 44 (1984), pp. 114–117; and "Modal Paradox: Parts and Counterparts, Points and Counterpoints," in French, Uehling, and Wettstein, eds., *Midwest Studies in Philosophy XI: Studies in Essentialism* (Minneapolis, Minn.: University of Minnesota Press, 1986), pp. 75–120. The last includes further bibliographical references.

modal logic. In this paper I shall present the objection(s) and my response. I shall also argue for the further claim (which I have not made elsewhere) that even the conventionally accepted system *B*, which is weaker than *S5* and independent of *S4*, has not been adequately justified as a fallacy-free system of reasoning about what might have been. The axioms characteristic of *B* are sentences of the form " $\phi \supset \Box \Diamond \phi$ ." That is, *B* is characterized by the principle that for any true proposition *p*, the proposition that *p* is possibly true is itself necessarily true. Here, however, I shall not argue for the strong claim (analogous to my claim in connection with *S4*) that some true proposition *p* is such that the proposition that *p* is necessarily possible is untrue. (I believe that the characteristic *B* principle may well have no such counterexamples.) I contend only that, even if the *B* axioms are in fact true, and even if they are necessarily true, it seems to be logically possible for some proposition *p* to be true while the proposition that *p* is necessarily possible is at the same time false. Thus, even if the *B* principle is necessarily true, its alleged status as a logical (or analytic) truth remains in need of justification. Similar arguments may be made against other proposed extensions of the weak modal system *T*. If I am correct, insofar as modal logic is concerned exclusively with the logic of metaphysical modality, and not also with other, nonlogical features of metaphysical modality, *T* may well be the one and only (strongest) correct system of (first-order) propositional modal logic.<sup>2</sup>

## I.

The case against *S4* modal logic stems from the intuition (which

---

<sup>2</sup>Metaphysical modal logic concerns metaphysical (or alethic) necessity and metaphysical (alethic) possibility, or necessity and possibility *tout court*—as opposed to such other types of modality as physical necessity, epistemic necessity, etc. The (strongest) correct system of logic for some other modality need not coincide with that for metaphysical modality. (The characteristic principle of *T* that any proposition that must be true is true must already fail in deontic modal logic, the logic of what is morally required to be the case and what is morally permitted to be the case.) Throughout this paper I am concerned primarily with metaphysical modality. Where I speak simply of "modal logic," the reader is to understand that only metaphysical modal logic is under discussion. My use of such modal locutions as "necessary," "might have," etc. is to be similarly construed throughout, unless otherwise indicated.

many of my opponents share) that a particular material artifact—say, a particular wooden table which we may call “Woody”—could have originated from matter slightly different from its actual original matter  $m^*$  (while retaining its numerical identity, or its *haecceity*) but not from entirely different matter. Wherever one may choose to draw the line between what matter Woody might have originated from and what matter Woody could not have originated from, it would seem that, by stretching things to the limit, we may select some (presumably scattered) matter  $m$  such that, although Woody could not have originated from  $m$ ,  $m$  is close enough to being a possibility for Woody that if Woody *had* originated from certain matter  $m'$  that *is in fact* possible for Woody—matter differing in as many molecules from the actual original matter  $m^*$  as possible, and sharing as many molecules with  $m$  as possible, while remaining a possibility for Woody—then it *would have been* possible for Woody to have originated from  $m$ , even though it is not *actually* possible. Even if one denies that there is a sharp line to be drawn between what matter is and what matter is not possible for the origin of Woody, by stretching things to whatever sort of limit remains (such as an interval of vagueness and indeterminacy in lieu of a dividing line between what is and is not possible), there will still be some matter  $m$  such that Woody (just barely) determinately could not have originated from  $m$ , yet the claim that this is itself necessary is untrue (or not “true to the maximum degree,” or whatever), and in addition, unfalse. Either way, the conditional claim (which is an axiom of  $S4$ ) that if Woody necessarily does not originate from  $m$ , then it is necessary that Woody necessarily does not thus originate fails. (It suffers the same truth-value status as its consequent.) Also failing is the inference from the antecedent of this conditional to the consequent, since the premise of the inference is altogether true and the conclusion is not.  $S4$  modal logic is fallacious.

I supplemented my argument against  $S4$  with a particular conception of what possible worlds are—in conjunction with the standard identification of necessity with truth in every possible world and possibility with truth in at least one possible world. As with many contemporary philosophers of modality, I conceive of possible worlds as certain sorts of (in some sense) maximal abstract entities *according to which* certain things (facts, states of affairs) obtain and certain other such things do not obtain. Possible worlds

are total *ways things might have been* (David Lewis). A possible world is something like a total history that might have obtained concerning everything in the cosmos (Saul Kripke), or a maximal property or state that the cosmos might have had or been in (Kripke, Robert Stalnaker), or a maximal state of affairs (Alvin Plantinga) or maximal scenario (myself) that might have obtained. For most purposes, one may conceive of a possible world as an infinitely long, complex, and detailed set of states of affairs or (potential) facts or statements (that is, an infinite set of structured propositions, more or less as Russell conceived propositions<sup>3</sup>), one that does not leave any question of fact undecided (Robert Adams). Since the actual world is itself a possible world, it too is conceived of as a maximal scenario or history, and may be conceived of as a maximally comprehensive set of statements, in this case the set of all statements that are in fact true.

More accurately, a possible world may be conceived of as a set of (potential) facts or statements that does not leave any of a very comprehensive range of questions of fact undecided. Some of the facts that are decided may in some cases determine that certain other statements are neither true nor false, owing to false presuppositions, category mistakes ("sortal incorrectness"), vagueness, or something else. If Frege was right, for example, the fact that there is no present King of France determines that the statement that the present King of France is bald is neither true nor false, so that neither this statement nor its negation is included in the set of statements corresponding to the actual world. More importantly, certain meta-facts (or facts about possible worlds and sets of facts) cannot be included in such a set for familiar reasons concerning cardinality problems, since there are at least as many such meta-facts as there are subsets of any given infinite set of facts, and these subsets outnumber the facts in the given set.<sup>4</sup> A possible world,

<sup>3</sup>I mean to exclude here the modal logician's conception of a proposition as a set of possible worlds (or equivalently, as a characteristic function from possible worlds to truth values). It is not a good idea to think of possible worlds as sets of propositions, and at the same time to think of propositions as sets of possible worlds. For more on my favored Russellian conception of propositions, see my *Frege's Puzzle* (Cambridge, Mass.: The MIT Press/Bradford Books, 1986).

<sup>4</sup>See Selmer Bringsjord, "Are There Set Theoretic Possible Worlds?" *Analysis* 45 (1985), p. 64; Christopher Menzel, "On Set Theoretic Possible Worlds," *Analysis* 46 (1986), pp. 68–72; and Patrick Grim, "On Sets and Worlds: A Reply to Menzel," *Analysis* 46 (1986), pp. 186–191.

then, may be thought of as a set of statements of a certain restricted but still very comprehensive sort.

Recall that it is (just barely) impossible for Woody the table to have originated from certain matter  $m$ . Woody cannot be in the state of originating from  $m$ . That is, originating from  $m$  is a state metaphysically unavailable to Woody; it is a way that Woody cannot be. But it is still a way for an individual to be. Likewise, there is a total way for all things in general to be—a “maximal” set of (potential) facts, if you will—according to which Woody originates from  $m$ . Let us call this maximal way for things to be “ $W$ .” Since Woody originates from  $m$  according to  $W$ , and Woody metaphysically cannot do so,  $W$  is a total way things cannot be. A total way things cannot be is a total way for things to be such that things cannot be that way, a state or history for everything in the universe such that everything in the universe cannot be in that state or have that history, a maximal state of affairs or scenario that cannot obtain. Total ways things cannot be are thus also “worlds,” or maximal ways for things to be. They are impossible worlds. In fact, although  $W$  is an impossible world, there is a possible world  $W'$  (assuming  $m$  was chosen carefully enough, and ignoring for the moment the prospect of vagueness and regions of indeterminacy) according to which Woody originates from the matter  $m'$  instead of its actual original matter  $m^*$ , and if  $W'$  had obtained (as indeed it might have),  $W$  would have been a way things might have been rather than a way things cannot be;  $W$  would have been possible instead of impossible. Although  $W$  is impossible relative to the actual world, it is possible relative to  $W'$ , which is itself possible relative to the actual world. Thus  $W$  is a possibly possible world. Other impossible worlds may be not even possibly possible, but only possibly possibly possible, and so on. The binary relation between (possible or impossible) worlds of relative possibility—the modal relation of *accessibility*—is not transitive.

What are the limits on the admissibility of possible and impossible worlds? None to speak of. Any degree of variation and recombination qualifies. Some ways for things to be are not even possibly possibly . . . possible, for any degree of nesting. A world according to which Nathan Salmon is Henry Kissinger is such a world, for example, as is a world according to which Nathan Salmon is a Visa credit card account with the Bank of America. Since they are ways-for-things-to-be of a certain sort (viz., such that

things necessarily cannot be that way, and necessarily necessarily cannot be that way, and so on), these too are “worlds.” As far as I can tell, worlds need not even be logically consistent. A world according to which there is both life on Mars and no life on Mars is a way things cannot be on logical grounds alone. Hence this too is a “world,” a way for things to be. The only restriction on *worlds*, as opposed to lesser ways for things to be, is that they must be (in some sense) *maximal* (total, comprehensive) ways for things to be; for every statement of fact, either it or its denial must obtain according to a world—*modulo* cases of nonbivalence arising from presupposition failure, vagueness, etc., and subject to cardinality constraints if the totality of facts comprising a world are to form a set.

## II.

The first part of the standard objection to this account is summed up by David Lewis as follows:

Say I: This is no defence [of the essentialist doctrine that a table could not have originated from entirely different matter], this is capitulation [to radical anti-essentialism]. In these questions of haecceitism and essence, by what right do we ignore worlds that are deemed inaccessible? Accessible or not, they're still worlds. We still believe in them. Why don't they count? (*On the Plurality of Worlds*, Oxford: Basil Blackwell, 1986, p. 246).

This part of the objection may be spelled out further: Intransitive accessibility relations are introduced into modal semantics for the purpose of interpreting various “real” or restricted types of modalities, such as nomological necessity. A proposition is *nomologically necessary* in an arbitrary possible world  $w$  if and only if it is true in every possible world in which all of the laws of nature in  $w$  are true. For convenience, we may say that a world  $w'$  is *accessible to*, or *nomologically possible relative to*, a world  $w$  if every natural law of  $w$  is true in  $w'$ . Then we may say more succinctly that a proposition is nomologically necessary with respect to a possible world  $w$  if and only if it is true in every possible world accessible to  $w$ . More restrictedly, perhaps, a proposition is *physically necessary* with respect to an arbitrary possible world  $w$  if and only if it is true in every possible world in which all of the laws of physics in  $w$  are

true. Other restricted modalities require alternative accessibility relations: a proposition is said to be *necessary*, in the restricted sense in question, with respect to an arbitrary possible world  $w$  if and only if it is true in every possible world of such-and-such a restricted sort—the restriction in question depending on some appropriate relation to  $w$ . Such restrictions yield failures of the characteristic *S4* principle that any “necessary” truth is necessarily necessary, and even of the characteristic *B* principle that any truth is necessarily possible. Suppose, for example, that  $w$  and  $w'$  are worlds so different in their natural constitution that although every natural law of  $w$  is true in  $w'$  (so that  $w'$  is nomologically possible relative to  $w$ ), some of these natural laws of  $w$  are not natural laws in  $w'$  but merely accidental generalizations, while certain other generalizations not even true in  $w$  are additional natural laws in  $w'$ . Then a natural law of  $w$  (which is automatically nomologically necessary in  $w$ ) that is not also a natural law of  $w'$  will not be true in every world nomologically possible relative to  $w'$ , and hence will not be nomologically necessarily nomologically necessary in  $w$ . Similarly, a proposition that is true in  $w$  but violates one of the additional natural laws of  $w'$  will not be nomologically necessarily nomologically possible in  $w$ . In this restricted scheme, accessibility between worlds is neither transitive nor symmetric. It remains reflexive, of course—as long as the natural laws of a given world are true in that world. The fundamental characteristic *T* principle that any “necessary” truth is true is thereby preserved.

By contrast, the objection goes, the hallmark of metaphysical (alethic) necessity or necessity *tout court*—its distinguishing characteristic—is that it is completely unrestricted. Metaphysical necessity and possibility is the limiting case of restricted necessity and possibility, the case with no restrictions whatsoever. A proposition is necessary in this unrestricted sense with respect to a possible world  $w$  if and only if it is true in absolutely every possible world whatsoever, no restrictions. By contrast with the case of restricted modalities, the objection continues, my conception of a *metaphysically impossible world* is incoherent. Any possible world is *possible* in the unrestricted, metaphysical sense. Since my account admits the existence of a world  $W$  in which Woody originates from  $m$ , even though I deem this world “inaccessible” to the actual world, I implicitly acknowledge (contrary to my explicit pronouncements) that it is *not* necessary in the relevant, metaphysical sense of “neces-



sary" that Woody does not originate from  $m$ . Indeed, by admitting possible worlds of unlimited variation and recombination, I simply abandon true metaphysical essentialism. By my lights, any property is attached to anything in some possible world or other. I am a closet radical anti-essentialist.

This part of the objection brings with it an oft-used defense of  $S5$  modal logic. In the metaphysical, unrestricted senses of "necessary" and "possible," the characteristic  $S5$  principle that any possible truth is necessarily possible may be easily proved. Suppose  $p$  is a possible truth, that is, a proposition true in at least one possible world  $w$ . Then relative to any possible world  $w'$ , without exception, there is at least one possible world in which  $p$  is true—namely,  $w$ . It follows (given our assumption that  $p$  is possible) that it is necessary that  $p$  is possible. For in the unrestricted sense of "possible," one possible world in which  $p$  is true is all that is required for  $p$  to be "possible" relative to any given world  $w'$ , with no further restriction as to what sort of world  $p$  is true in or how that world is related to  $w'$ . There are similar direct proofs of the characteristic  $B$  and  $S4$  principles.

There remains my claim that such a world as  $W$ , in which Woody originates from  $m$ , is inaccessible to the actual world. The first part of the objection more or less ignores this claim as irrelevant, a red herring. The second part of the objection focuses on this claim. When such restricted modalities as nomological necessity or physical necessity are under discussion, the phrase "possible relative to" has a tolerably clear sense (given that we have a prior understanding of such notions as *law of nature* and *law of physics*). Such notions of accessibility are more or less sharply defined. My notion of *necessity* is also some restricted notion, since I deem some worlds inaccessible to others. Yet, the objection goes, I have not defined the restriction; I leave my use of the phrase "possible relative to" with no tolerably clear sense. It does not seem to mean much of anything; it is simply an *ad hoc* device for sweeping a serious difficulty under the rug. To quote Lewis again:

[W]e look in vain, in . . . many . . . places, for an account of what it means to deny that some world is 'relatively possible'. I think it is like saying: there are things such that, ignoring them, there are no such things. Ignoring all the worlds where such-and-such obnoxious things happen, it is impossible that such things happen. Yes. Small comfort (ibid., p. 248).

III.

The objection presented in the preceding section confuses or conflates two notions that must be kept sharply distinct: the generic notion of a way for things to be and the peculiarly modal notion of a way things *might have* been. Confusion between these two notions probably stems from an analogous ambiguity in the phrase “possible world.” The layman speaks of a “world” almost exclusively as a planet, though sometimes as the whole physical universe of atoms, molecules, planets, stars, galaxies, super-clusters, and what-have-you. By contrast, in the metaphysics of modality a *world* is an abstract entity according to which some things obtain and other things do not, such that all (or sort of all) such questions of fact are answered one way or the other. Modal worlds are not physical universes but intensional entities that represent things as being one way or another. Even Lewis, who in his metaphysical constructions idiosyncratically maintains the layman’s conception of a *world* as a whole physical universe, combines this conception with the metaphysician’s conception of a world as an entity *according to which* some states of affairs obtain (including, for Lewis, states of affairs concerning things not part of that world) and other such states of affairs do not, such that all (or sort of all) such questions of fact are answered by the “world.”<sup>5</sup> It is awkward to call these things simply “worlds,” since that term is so highly suggestive of the layman’s notion. Fortunately (or rather unfortunately!) Leibniz provided a more descriptive term: “possible world.”

There are two problems with this bit of Leibnizian terminology. The first problem concerns what the word “possible,” as it occurs in the phrase “possible world,” does not mean. In metaphysics when we call something a possible such-and-such, we generally mean that it is a such-and-such that might have existed, even if it does not. But whether or not possible worlds actually exist, in calling something a “possible world” most of us do not mean a world (*qua* total way for things to be, or maximal entity according to which some states of affairs obtain and others do not) that might

<sup>5</sup>I criticize Lewis’s views concerning the nature of possible worlds in my review of his *On the Plurality of Worlds*, in *The Philosophical Review* 97 (1988), pp. 237–244.

have existed, even if it does not. To think that the concept of a possible world is that of a world that might have existed is to misunderstand the function of the word “possible” in the phrase “possible world.”<sup>6</sup>

The second problem with the phrase “possible world” concerns what the word “possible” does mean there. For it means *something* there. Strictly speaking, a possible world is not a way for things to be that might have existed; it is a way for things to be such that things might have been that way. Similarly, a possible history or possible state for an individual is not a history or state that might have existed, but a history or state that the individual might have had or might have been in. Thus the word “possible” contributes some special meaning to the phrase, and more meaning than is accommodated by the generic notion of a total way-for-things-to-be-even-if-things-could-not-have-been-that-way. Strictly speaking, a possible world is not any old total way for things to be, but a modally special kind of total way for things to be, namely a total way that things *might have* been. A possible world is a total way for things to be that conforms to metaphysical constraints concerning what might have been. The generic notion of a total way for things to be is a notion without a proper term of its own. Aesthetic considerations aside, rather than let the phrase “possible world” do double duty for this generic notion as well as for the modal notion, we would be better off reserving it exclusively for the modal notion—for which it is certainly more apt—and using my highfalutin hyphenated phrase “total way-for-things-to-be-even-if-things-could-not-have-been-that-way” for the generic notion, or my modally unadorned phrase “total way for things to be,” or if worse comes to worst, the simple unadorned word “world.” In the best of all possible worlds, total ways for things to be are not called “possible worlds,” unless they are total ways things might have been.

---

<sup>6</sup>The objection of the preceding section need not depend in any way on this common misconstrual of the phrase “possible world,” although it probably often does. One who misunderstands the phrase “possible world” to mean *world that might have existed* will conclude that “impossible worlds” cannot exist. Possible worlds would emerge as the only worlds there could be, so that a (possible) thing is a world if and only if it is a “possible world.” It seems likely that this fallacy lies behind the common confusion of the generic notion of a world and the modal notion of a possible world.

Whatever the source of the confusion between the generic notion of a way for things to be and the modal notion of a way things might have been, this confusion is very probably the primary source of the idea that metaphysical modality is the limiting case of restricted modalities, that metaphysical necessity and possibility is the unrestricted, and hence the least restricted, type of necessity and possibility. For metaphysical necessity is indeed truth in all ways things might have been (modal, not generic), and metaphysical possibility is indeed truth in at least one way things might have been (modal, not generic).

Metaphysical modality is definitely *not* an unrestricted limiting case. There are more modalities in Plato's heaven than are dreamt of in my critics' philosophy, and some of these are even less restrictive than metaphysical modality. One less restrictive type of modality is provided by *mathematical necessity* and *mathematical possibility*. A proposition is mathematically necessary if its truth is required by the laws of mathematics alone, and mathematically possible if its truth is not precluded by the laws of mathematics alone. Many metaphysical impossibilities are mathematically possible, for example, Nathan Salmon being a Visa credit card account with the Bank of America. Another type of modality less restrictive than metaphysical modality is provided by what is sometimes called "logical necessity" and "logical possibility," to be distinguished from genuinely metaphysical necessity and possibility, or necessity and possibility *tout court*. A proposition is logically necessary if its truth is required on logical grounds alone, logically possible if its truth is not ruled out by logic alone (that is, if its negation is not logically necessary). Thus whereas it is logically necessary that Nathan Salmon is not somebody other than Nathan Salmon, and it is also logically necessary that either Nathan Salmon is a Visa credit card account with the Bank of America or he is not, it is not logically necessary that Nathan Salmon is not a credit card account. Although there is a way things logically could be according to which I am a credit card account, there is no way things metaphysically might have been according to which I am a credit card account. This illustrates the restricted nature of metaphysical modality. Some logically possible worlds must be "ignored." Metaphysical necessity is truth in every logically possible world of a certain restricted sort.

What is the restriction? To worlds that are metaphysically pos-

sible. (What else!) When we identify necessity with truth in every possible world, the word “possible” means *something* there, and what it means there places a restriction on the sort of worlds under consideration. The metaphysical notion of possibility restricts the logical notion of possibility, in a manner exactly analogous to that in which the notion of *natural law* involved in the notion of nomological necessity restricts the metaphysical notion of possibility. Just as nomological possibility is a special kind of metaphysical possibility, so metaphysical possibility is a special kind of logical possibility.<sup>7</sup>

Even logical necessity may be seen as observing some restriction: a proposition is logically necessary (with respect to a world  $w$ ) if and only if it is true in every *logically consistent* world (according to  $w$ ), whether metaphysically possible or not—or every world in which the laws and rules of logic (in  $w$ ) obtain (including the logical prohibition on inconsistency).<sup>8</sup> The logically inconsistent worlds do not count as regards what is logically necessary. Still, logical modality is considerably freer of restriction than metaphysical modality. With its freedom from the additional constraint of metaphysical possibility, logical necessity may be construed as accommodating all of the axioms and rules of *S5*. But if logical modality is unrestrictive enough to accommodate all of the axioms and rules

---

<sup>7</sup>Timothy Williamson has pointed out that this may be strictly false, since (as David Kaplan has shown) there are sentences that are valid in the logic of indexicals and that do not express metaphysically necessary truths, for example “If Saul Kripke is an anthropologist, then Saul Kripke is actually an anthropologist.” I believe, however, that insofar as propositions (as opposed to sentences) may be appropriately called “logically valid” or “not logically valid,” the propositions expressed by such sentences are not logically valid even though the sentences themselves are. (Conversely, some sentences that are not logically valid express propositions that are, for example, “All bachelors are unmarried men.”) Cf. *Frege’s Puzzle*, pp. 132–151, and especially p. 177, note 1. The important point here is that some logically possible (that is, consistent) propositions are nevertheless metaphysically impossible.

<sup>8</sup>If  $w$  is itself logically consistent, this rules out worlds in which such logical truths as the Law of Noncontradiction do not obtain. What about an inconsistent world according to which there is both life on Mars and no life on Mars and yet (by logic) no proposition and its negation are both true? (I owe this marvelous example to Saul Kripke, who has used it for a different but related purpose.) This had better count somehow as a world in which the Law of Noncontradiction does *not* obtain, in the relevant sense. Otherwise, such contradictions will emerge as logical possibilities.

of *S5*, it may not be restrictive enough to zero in on *S5*. Depending on what counts as logically possible, the interpretation of the diamond “ $\diamond$ ” as logical possibility instead of metaphysical possibility could turn ‘ $\diamond\phi$ ’ into a logical truth for every logically consistent formula  $\phi$ . It would then become a logical truth that Woody “might have” originated from *m*, and that Nathan Salmon “might have” been a credit card account. Even if we essentialists are wrong and metaphysical necessity does not extend beyond logical necessity, the logic of logical necessity can extend far beyond that of metaphysical necessity.<sup>9</sup>

If worlds include ways things metaphysically cannot be in addition to ways things metaphysically might have been, then the idea that metaphysical necessity corresponds to truth in every world whatsoever is flatly mistaken. If worlds include ways things logically cannot be, then no proposition is true according to every world and every proposition is true according to some world. I know of no standard or conventional sense of “possible” on which even the proposition that Nathan Salmon is somebody other than Nathan Salmon is “possible.” It is not clear that there would be any interest, other than purely formal interest, in a completely unrestricted notion of modality on which anything is possible and nothing is necessary—and there is not much purely formal in-

---

<sup>9</sup>Thus whereas it is metaphysically impossible on my view for Woody to originate from *m*, it may nevertheless be logically true (and hence logically necessary) that it is logically possible that Woody so originates. Whether the sentence “It is logically possible that Woody originated from *m*” should itself count as a logical truth may depend on whether logical necessity and possibility are treated as attributes of sentences, or rather as attributes of propositional contents. See note 7 above. It is arguable that the logical (as opposed to metaphysical) possibility of truth for the proposition that Woody originated from *m* is itself a truth of pure logic. Alternatively, if logical possibility is an attribute of sentences rather than of their contents, it is arguable that the logic of logical necessity and possibility should take into consideration the logical possibility of the sentence “Woody originates from *m*” being analytically false while retaining its logical form (expressing, for example, the proposition that Venus is distinct from Venus). In that case, it need not be a truth of logic (although it would still be true) that “Woody originated from *m*” is logically possible. Even under this construal, however, *S5* may not be the appropriate (first-order) propositional logic of logical necessity. The rule of necessitation (which licenses the inference of ‘ $\Box\phi$ ’ from a subsidiary proof of  $\phi$ ) is inapplicable to such logical validities as “If Saul Kripke is an anthropologist, then Saul Kripke is actually an anthropologist.” (See note 17 below.)

terest in this unrestricted notion. Such a notion would preserve the characteristic *S4* axiom schema, but perhaps at the cost of turning ' $\diamond\phi$ ' into a logical truth for every formula  $\phi$ , and thereby ruling out the inference rule of necessitation (which licenses the inference from a logical theorem  $\phi$  to ' $\Box\phi$ ') as well as the characteristic axiom schema of *B* and hence also that of *S5*. (The last, in fact, would be replaced by its negation.) Even if there is interest in such a notion, it has nothing to do with metaphysical modality. Surely it is metaphysically impossible that there should be life on Mars and no life on Mars at the same time. The failure of the characteristic *B* axiom schema in the case of the completely unrestricted interpretation of the modal operators demonstrates that there must be some fallacy in the "proof," presented in the preceding section, that unrestricted modality honors *S5*.

Do worlds, *qua* ways for things to be, include ways things cannot be in addition to ways things might have been? I know of no plausible grounds for denying that they do. Indeed, nearly any plausible argument for the existence of ways things might have been (including those arguments offered by my opponents)<sup>10</sup> affords an analogous and parallel argument for ways things cannot be, even ways things cannot be on logical grounds alone. Every argument I am aware of against impossible worlds in favor of only possible worlds confuses ways for things to be with ways things might have been, or worse, confuses ways things cannot be with ways for things to be that cannot exist—or worse yet, commits both errors. The fact that Woody cannot originate from *m* entails that originating from *m* is a way Woody cannot be. It follows from the latter that Woody originating from *m* and Socrates being wise and . . . (where "all" questions of fact are fixed), is a maximal way that things (in general) cannot be. It follows from the fact that Woody cannot originate from *m*, therefore, that there is a maximal way things cannot be. Likewise, it follows from the fact that I cannot be somebody other than myself, that me being somebody other than myself and Socrates being wise, etc., is also a way things cannot be. We should not resist these inferences; we should draw them, and see where they lead. At the very least we should refrain from as-

---

<sup>10</sup>Cf. "Impossible Worlds," pp. 116–117; and Margery Bedford Naylor, "A Note on David Lewis's Realism About Possible Worlds," *Analysis* 46 (1986), pp. 28–29.

serting their premises while rejecting their conclusions, since they are valid.

An impossible world like  $W$  may be seen as merely a variation of a genuinely possible world. Consider the “maximal” set of statements that would have been true if  $m$  had been formed into a table and Woody had never been constructed at all. Let us call the (possible) table that would have been formed from  $m$  if  $m$  had formed a table “Mia,” and let us call this set of statements “ $K_{Mia}$ .” Now there is surely a “maximal,” coherent set of statements  $K$  like  $K_{Mia}$  except that every statement in  $K_{Mia}$  concerning Mia (or concerning the table formed from  $m$ ) is replaced by the corresponding statement concerning Woody, and every statement concerning Woody is replaced by the corresponding statement concerning Mia (or the table that actually would have been formed from  $m$ ), with whatever further additions and deletions are required by these changes. The world  $W$  is simply the way-for-things-to-be determined by  $K$ . Indeed,  $W$  is just like the possible world  $W_{Mia}$  corresponding to  $K_{Mia}$  (the maximal scenario that would have obtained if  $m$  had been formed into a table and Woody had never been constructed), except for the substitution of certain “components” (nonmaximal scenarios, as it were). Since  $W$  is a world according to which Woody originates from  $m$ , and by hypothesis Woody cannot thus originate, we have here what so many philosophers have so often repudiated: an impossible world. But what is there to repudiate? World  $W$  is just the maximal way-for-things-to-be corresponding to a particular set of statements or (potential) facts, something of the same ontological category or sort as the genuinely possible world  $W_{Mia}$ . The key difference between  $W_{Mia}$  and  $W$  is modal rather than ontological-categorical. The former might have been realized whereas the latter could not have been realized; the former is a way things might have been whereas the latter is a way things could not have been. Both are ways for things to be, and in that sense, ontologically on a par.<sup>11</sup>

---

<sup>11</sup>If anything,  $W_{Mia}$  is the more dubious of the two, since it directly involves Mia, which does not actually exist, in place of Woody, which actually exists. But let us not worry about this potentially significant ontological difference here. If the truth be told, my own view is that most of the worlds quantified over by modal semanticists do not actually exist, though they might have existed, or possibly might have existed, or possibly possibly might have existed, etc. I do not see this as a decisive reason



## IV.

Given this conception of metaphysically possible worlds as forming a restricted subclass of more things of the same ontological category, one cannot rely on the mere existence or nonexistence of worlds according to which it is the case that such-and-such in order to determine whether such-and-such is possible or impossible. It is metaphysically impossible for Woody to originate from  $m$ , yet there are many worlds according to which Woody so originates. On my conception, the notions of metaphysical necessity and possibility are not defined or analyzed in terms of the apparatus of possible worlds. The order of analysis is just the reverse: a possible world is understood to be a total way things might have been (or a maximal scenario that might have obtained, etc.), relying on one's prior understanding of the modal notion of what *might have been*. What is possible and what is impossible according to a world is determined by the world itself. Recall that worlds are *maximal* or *total* ways for things to be, deciding all (or a very comprehensive class of) questions of fact. They are not silent concerning all questions of modal fact, since these too are questions of fact. If  $p$  is a nonmodal proposition, then one (partial) way for things to be is for  $p$  to be a necessary truth, and another is for not- $p$  to be possible. Among the facts (or statements of fact, etc.) that comprise (or obtain according to) a world are such peculiarly modal facts, facts of the form "It is necessary that such-and-such" or "It is possible that such-and-such." It is a fact of the actual world, for example, that it is necessary that Woody does not originate from  $m$ , and this fact is included among the facts that comprise the actual world. Given this conception of what a world is, the relevant notion of relative possibility, or accessibility, is perfectly straightforward. If a definition is wanted, it is this: a world  $w'$  is *metaphysically possible relative to* a world  $w$  if and only if every fact of  $w'$  is a possibility in  $w$  (that is every proposition that is true according to  $w'$  is possible according to  $w$ ). Equivalently,  $w'$  is metaphysically possible relative to  $w$  if and only if every necessary fact of  $w$  obtains in  $w'$  (that is, every proposition that is necessary ac-

---

not to quantify over them, as long as one keeps one's ontology straight. Cf. my "Existence," in J. Tomberlin, ed., *Philosophical Perspectives I: Metaphysics* (Atascadero, Calif.: Ridgeview, 1987), pp. 49–108.

cording to  $w$  is true according to  $w'$ ). If we assume that one question of fact decided by any maximal scenario (or total way for things to be) is the question of whether a given alternative maximal scenario is a scenario that might have obtained (a way things might have been), and we note that on every consistent maximal scenario it itself is the only maximal scenario that obtains, we may prove that every necessary fact of a consistent maximal scenario  $w$  obtains in a given alternative maximal scenario  $w'$  if and only if on scenario  $w$ ,  $w'$  is a scenario that might have obtained. (If "maximal" scenarios are sets of such things as purported facts, then such facts as that an alternative maximal scenario is a maximal scenario that might have obtained will be meta-facts, which obtain according to the given set of facts not by being included directly as elements of the set but only implicitly by virtue of the facts that are included in the set.) If we confine our attention to consistent maximal scenarios, we may thus put our "definition" another way: to say that a maximal scenario (or total way for things to be)  $w'$  is *metaphysically possible relative to* a consistent maximal scenario  $w$  is to say that on scenario  $w$ ,  $w'$  is a scenario that might have obtained (a way things might have been). More simply, a world  $w'$  is accessible to a consistent world  $w$  if and only if  $w'$  is possible in  $w$ . Being "accessible to" or "possible relative to" a consistent world is simply being possible *according to* that world, nothing more and nothing less. On this conception, what is possible and what is necessary at a given world is not imposed from above by a mysterious and unanalyzed accessibility relation among worlds; rather, a world's accessibility relations to other worlds is internal to the world, *via* the possibilities at that world.<sup>12</sup>

---

<sup>12</sup>Lewis's complaint that "we look in vain, in . . . many . . . places, for an account of what it means to deny that some world is 'relatively possible' " is unjustified. The definition I propose here of the accessibility relation is the natural one, and as Saul Kripke pointed out to me, it follows precisely the characterization of accessibility that he had offered originally in "Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi," *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 9 (1963), pp. 67–96, at p. 70; and again in "Semantical Considerations on Modal Logic," in L. Linsky, ed., *Reference and Modality* (New York, N.Y.: Oxford University Press, 1971), pp. 63–72, at p. 64. There is no suggestion in these pioneering works that such subsystems as  $T$ ,  $B$ , or  $S4$  arise from special restrictions on metaphysical modality; instead accessibility is explained in terms of propositions being (metaphysically) possible in worlds. Kripke has informed me (in discussion and personal correspondence) that

It follows, given this conception, that a proposition is metaphysically necessary according to a consistent world  $w$  if and only if it is true in every world metaphysically possible relative to  $w$ , and a proposition is metaphysically possible according to a consistent world  $w$  if and only if it is true in at least one world metaphysically possible relative to  $w$ . These are not definitions of metaphysical necessity and possibility. They are theorems that follow from the definition of relative possibility. One must have a prior understanding of metaphysical modality in order to grasp the notion of it being the case that everything that *must* be so on one scenario is so on another scenario (the notion of one world being *possible relative to* another)—as well as the closely related notion of it being the case on one scenario that another scenario is a scenario that *might have* obtained (the notion of one world being *possible according to* another). The idea that the notion of a possible world comes first, and explains the notion of metaphysical modality, is of a piece with the same mythology that gave us the idea that metaphysical necessity is truth in every world whatsoever, without restriction. The notion of metaphysical modality comes first, and like every notion of modality, it is restricted.

There is one alternative yet to be considered. One may choose to ignore ways things could not have been, confining one's sights always and without exception to ways things actually might have been. One may stipulate that a proposition is necessary with respect to an arbitrary possible world  $w$  if and only if it is true in every world accessible to the actual world—never mind worlds accessible to  $w$ —and likewise that a proposition is possible with re-

---

he is sympathetic to many of the positions advanced in this paper, having seriously considered whether the conventional presupposition that the basic modal logic is  $S5$  is justified. He now believes he should have stressed both that his use of an accessibility relation does not make "possible" (as applied to worlds) into a dyadic predicate any more than the natural treatment of baldness in possible-world discourse as a binary relation between individuals and worlds makes "is bald" into a dyadic predicate, and that unless we have  $S4$ , strictly speaking, many of the worlds are not "possible," but only "possibly possible," and so on. Whereas Kripke shares some of my controversial views concerning the logic of metaphysical modality, he is not fully convinced that  $S4$  modal logic is invalidated in cases like that involving Woody and  $m$  (though he tells me he is nearly convinced). Cf. *Naming and Necessity*, p. 51n. See also *Reference and Essence*, pp. 240–252; and "Modal Paradox: Parts and Counterparts, Points and Counterpoints," especially pp. 89–95.

spect to an arbitrary possible world  $w$  if and only if it is true in at least one world accessible to the actual world. One may accordingly declare it impossible that Woody even *might have* originated from  $m$ , since one is ignoring possibly possible but impossible worlds like  $W$ , worlds that are once removed from the actual world on the scale of accessibility and in which Woody originates from  $m$ . One may then ignore accessibility altogether. We have finally zeroed in on S5 modal logic.

This is the ostrich approach to metaphysical modality. It is not a very happy alternative.<sup>13</sup> The ostrich approach flies in the face of the very meanings of the words “necessary” and “possible.” On any standard or conventional sense of “possible” in English, a sentence of the form “It is possible that such-and-such” is true if there is a possible (in the same sense) scenario, a way things might have been, according to which it is the case that such-and-such. Certainly this is so with respect to the metaphysical sense of “possible.” Likewise, in English, it is simply incorrect to say “It is necessary that such-and-such” when there is a possible scenario according to which it is not the case that such-and-such. In particular, therefore, as long as there is a possible scenario according to which it is possible for Woody to have originated from  $m$ , it is true (in English) to say “It is possible that it is possible that Woody originates from  $m$ ,” and one cannot correctly say (in English) “It is necessary that it is necessary that Woody does not originate from  $m$ ” (or “It is impossible that Woody *might have* originated from  $m$ ”). If the possible scenarios (such as  $W'$ ) that verify a possibility claim or falsify a necessity claim draw our attention to inaccessible worlds, then we are obliged to pay attention to those inaccessible worlds. We ignore them to our own detriment, counting what is true false and what is false true.

Surprisingly, the ostrich approach has nevertheless ascended to the status of orthodoxy. It is precisely the approach followed by my critics. The most obvious sign of the ostrich approach is the explicit denial of impossible worlds, but there are a number of additional signs, several of which manifest themselves in the objection presented in Section II above. If one ignores impossible worlds altogether, then ways things might have been are the only

---

<sup>13</sup>In *Reference and Essence* I referred (p. 239) to this philosophical position as “a narrow-minded form of modal ethnocentrism.”

ways for things to be that are left. The distinction between the generic notion and the modal notion loses all significance. If one confines one's sights to genuinely possible worlds, disavowing the impossible worlds, then metaphysical modality emerges as the limiting case—the “unrestricted” modality that takes account of “every” world—and *S5* emerges as its proper logic. Metaphysical modality appears unrestricted because the restriction to metaphysically possible worlds is already built into one's practice concerning which worlds to pay attention to and to quantify over. If certain entities are ignored entirely and always, then they are not even seen as things that are ignored. Since there is no possible world in which Woody originates from *m*, and possible worlds are the only worlds taken into consideration, one will insist that it is necessary that it is necessary that Woody does not originate from *m*, and that it is necessary that it is necessary that it is necessary . . . that Woody does not so originate, with as many iterations as one pleases. If some iconoclast comes along and argues that some worlds are inaccessible and that in some of these Woody originates from *m*, those who ignore impossible worlds altogether will be puzzled as to what this philosopher could possibly mean by “inaccessible,” and hence by “possible” and “might have.” Whatever restricted sort of modality the modal iconoclast means by these terms, it would seem to be based on some completely unexplained restriction among the *possible* worlds, for these are the only worlds that are ever considered. When the modal iconoclast protests that in pleading for inaccessible worlds he is not talking about a special and peculiar sort of possible world but about worlds of a sort entirely ignored by the friends of *S5*, those who ignore these worlds will shrug and dismiss these protests as lacking in substance. For in restricting their quantifications over worlds always to possible worlds, they can hardly help but misconstrue the modal iconoclast's claims concerning worlds in general, misinterpreting them as puzzling claims concerning possible worlds. Since he maintains that there are worlds in which Woody does indeed originate from *m*, the modal iconoclast is seen by those who quantify over only possible worlds as capitulating to anti-essentialism. Any such worlds would have to be possible, no matter what the modal iconoclast may mean by calling them “inaccessible,” since no other type of world is ever recognized and quantified over, no matter what anyone says. The situation is not unlike that of a philosopher who

tries to persuade a pure set theorist, whose quantifiers range only over sets, of the existence of ur-elements (non-set elements), and who is misunderstood as rejecting Extensionality by postulating a plurality of empty sets.

The practice in modal semantics of ignoring worlds that are not possible according to the actual world leads theorists into understanding something different with the use of our terms “necessary” and “possible” from what they mean in English. Specifically, the ostrich approach misconstrues the simple modal term “necessary” to mean the modally complex concept of *actual* necessity, or necessity according to  $W_{@}$ , where  $W_{@}$  is the actual world. Likewise, the ostrich approach misconstrues “possible” to mean *actual* possibility, or possibility according to  $W_{@}$ . The simple modal concepts of necessity and possibility *simpliciter*—the real meanings of the simple modal terms “necessary” and “possible”—are not the same as the concepts of actual necessity and actual possibility, necessity and possibility according to the actual world. In exactly the same way, the concept of a philosopher is not the same as that of an actual philosopher. The difference shows up in modal contexts. Whereas it was not necessary for Saul Kripke to have been a philosopher, he actually is a philosopher and hence (in the indexical sense of “actually”) it is necessary that he be *actually* a philosopher—since in every possible world, the actual world (indexical sense again) is one in which he is a philosopher. Likewise, whereas it is not necessary that it be necessary that Woody not originate from  $m$ , it is actually necessary that Woody does not so originate, and hence it is necessary that it be *actually* necessary that Woody does not so originate. In effect, the ostrich approach prevents us from speaking of nested modalities altogether, instructing us to misconstrue iterations of modal operators in our speech as redundant embellishments that make no significant contribution to cognitive information content, as mere stuttering. But ignoring impossible worlds does not make them go away, and reinterpreting someone’s words to mean what they do not in fact mean does not make the actual meaning go away. Although Woody’s originating from  $m$  is impossible, the presence of worlds such as  $W$ , in which Woody originates from  $m$  (and hence, which are impossible) but which are possible according to some possible worlds, makes *something* true as regards the prospect of Woody’s so originating. This *something* is expressed in English by saying that the prospect in question is

“possibly possible.” The *S5* theorist’s misconstrual of English makes nested modality unseen, but it does not make nested modality vanish. The modal iconoclast may echo the words of his colleague: In these questions of haecceitism and essence, by what right do we ignore worlds that are inaccessible? Accessible or not, they’re still worlds. Why don’t they count? Ignoring all the possibly possible worlds where such-and-such obnoxious things happen, it is impossible that such things even *might* happen. Yes. Small comfort.

The ostrich approach may offer comfort of sorts, but certainly no illumination. It is not I who ignore inaccessible worlds. I acknowledge them and give them their full due, no more and no less. It is my critic, the friend of *S5*, who ignores them altogether. In pleading for inaccessible worlds, I am not drawing an unexplained distinction among the worlds that my opponents recognize, and proposing to ignore those on one side of the undefined boundary line. I am calling attention to worlds to which my opponents pay no attention (other than to repudiate).

## V.

The world *W*, in which Woody originates from *m*, is a way things could not have been. Nevertheless, there is a way things might have been, *W'*, in which Woody originates from *m'* instead of from *m\**, and in (according to, relative to, from the point of view of) *W'*, *W* is a way things might have been, as is the way things actually are. The denial of this is highly counterintuitive.<sup>14</sup> The impossible world *W* is thus only contingently impossible. No doubt it is an essential property of any way things could not have been that it is a way for things to be. And of course, some impossible worlds (such as a world according to which I am a credit card account) are es-

---

<sup>14</sup>I am ignoring here the complications introduced by indeterminacies and regions of vagueness. These complications complicate, but do not significantly alter, the points I am making. Roughly, the idea is that it may in some cases be neither true nor false according to a world *w* (owing to vagueness in the notion of metaphysical necessity) whether a certain fact obtaining in *w* is necessary. This, in turn, would inject some indeterminacy into the accessibility relation, so that some worlds may be neither definitely possible nor definitely impossible relative to others. These complications are discussed in some detail in “Modal Paradox: Parts and Counterparts, Points and Counterpoints.”

entially impossible. But others are not. Similarly, it is only a contingent fact about  $W'$  that it is a way things might have been rather than a way things could not have been. For there is some matter  $m''$  that Woody might have originated from in lieu of  $m^*$ , and that differs considerably enough from  $m'$  (though overlapping just enough with the actual original matter  $m^*$  to remain a possibility for Woody's origin) that if Woody had originated from  $m''$ , it would then have been impossible for Woody to have originated from  $m'$ .<sup>15</sup> Let  $W''$  be a possible world in which Woody originates from  $m''$ . From the point of view of  $W''$ ,  $W'$  is impossible. Perhaps the actual world is essentially possible. (That is, it may be that the actual world is possible relative to every world possible relative to it.) Even so, some possible worlds are like  $W'$ , only contingently possible. Whether a world is possible or not can be a question of contingent fact like any other question of contingent fact.

This sort of consideration uncovers the fallacy in the "proof," presented in Section II, of the characteristic S5 principle that any possible truth is necessarily possible. The argument was that if a proposition  $p$  is true in some possible world  $w$ , then no matter what possible world one considers, from its point of view  $p$  is true in at least one possible world, namely  $w$ , so that in the metaphysical sense of "possible" (in which one possible world in which  $p$  is true is all that is required for  $p$  to be possible with respect to any given world), if  $p$  is possible it is necessarily possible. This argument is framed with an ambiguous usage of the phrase "possible world," indiscriminately meaning either a way for things to be or a way things might have been. The argument is therefore susceptible to

---

<sup>15</sup>As long as some overlap is required and total replacement prohibited, such matter is always possible. Since  $m'$  is a possibility for Woody, there will be some overlap between  $m^*$  and  $m'$ . Simply replace as much of  $m^*$ 's overlap with  $m'$  as allowable with completely new matter, while preserving the remainder of  $m^*$ , including the entire portion of  $m^*$  replaced in  $m'$ . The resulting matter is  $m''$ . It differs from  $m'$  by more than the difference between  $m^*$  and  $m'$ , since it fully restores all of  $m^*$ 's matter that was replaced in  $m'$ —it duplicates the entire difference between  $m^*$  and  $m'$ —and in addition replaces some of the remaining matter of  $m'$  with new matter. If the proportion of required overlap is more than one-half (as seems reasonable), some overlap between  $m''$  and  $m'$  will remain, but not enough. Since the matter in  $m^*$  that was replaced in  $m'$  has been restored in full, and the maximal replacement by new matter is effected entirely elsewhere in  $m'$ , the resulting matter  $m''$  exceeds the allowable nonoverlap with  $m'$  by exactly the restored matter of  $m^*$ .



two conflicting interpretations. Since our concern is with the logic of what might have been, the argument is of considerably greater philosophical significance when it is interpreted as concerning genuinely possible worlds, rather than worlds in general. Under this interpretation the argument fallaciously presupposes that worlds that are possible in the actual here-and-now are also possible even according to alternative possible worlds.<sup>16</sup> This assumption, though perhaps understandable given the common confusion between possible worlds and worlds in general, is intuitively incorrect. The standard “proofs” of the characteristic *B* and *S4* principles likewise involve equivocation between the generic and properly modal sense of the phrase “possible world,” resulting in fallacious presuppositions concerning the essentiality of the property of being a possible world (*B*) or that of not being a possible world (*S4*).

Believers in *S5* as a correct system of reasoning (in propositional logic) about what might have been must claim that it is an essential property of any way things might have been that things might have been that way. Similarly, believers in the weaker *S4* modal logic (and hence also believers in *S5*) must claim that it is an essential property of any way things could not have been that things could not have been that way. Believers in *B* modal logic (and hence also believers in *S5*) must claim that it is an essential property of the way things actually are that things might have been that way. These claims are versions of essentialism. They are doctrines to the effect that certain properties (in this case, certain modal properties) of certain sorts of things (possible worlds, impossible worlds, and the actual world, respectively) are properties that these things could not fail to have. More than this, since their claim is that *S5*, *S4*, or *B* is a correct *logic* of what might have been and of

---

<sup>16</sup>When the purported “proof” of the characteristic *S5* principle is interpreted instead (less interestingly) as concerning all worlds without exception, whether genuinely possible or not, it commits a similar error. Under this interpretation the argument fallaciously presupposes that worlds that are available in the actual here-and-now as ones in which a given proposition is true remain available as such even according to alternative impossible worlds. Let *w* be a world in which a given proposition *p* is true. One cannot correctly conclude that no matter what world one considers, possible or not, *w* is still one world in which *p* is true. There are radically impossible worlds according to which *p* is not true in *w* or in any other world.

what must be, the essentialism espoused must be held to be not merely metaphysically true but true by the very *logic* of (metaphysical) necessity and possibility. The essentialism must be held to be not the metaphysically substantive sort of essentialism that requires Woody not to originate from *m* and me not to be a credit card account, but the minimal, vacuous, and trivial sort of essentialism that requires Woody to be such as to originate or not originate from *m*, that requires me not to be somebody other than myself, that requires Mars not to be such as to contain life and not to contain life at the same time. This does not weaken the import of the essentialist claims. On the contrary, the logical nature of the claims makes them extremely strong versions of essentialism. The claim is not merely that such-and-such worlds are essentially thus-and-so, but that they are essentially thus-and-so *by logic alone*. It is not merely by virtue of the laws of metaphysics that these worlds are supposed to be essentially thus-and-so, but by virtue of the very laws of logic and nothing more. The doctrine that some properties of some things are properties that on logical grounds alone these things could not fail to have is by itself the most trivial type of essentialism—because it is entirely nonspecific. The doctrine that *such-and-such* properties of *so-and-so* things are properties that on logical grounds alone these things could not fail to have is a horse of a different color. The logical essentialism concerning worlds that the friends of *S5*, *S4*, and *B* are committed to is some seriously committed essentialism. It is essentialism of the most committed type.

In fact, the logical essentialism concerning worlds that the friends of stronger modal logics are committed to seems intuitively false. At the very least, it requires substantial justification. The possible world *W'* is a way things are not but might have been; it is a way-for-things-to-be that is not realized, but might have been realized. This is just to say that it is a contingent or accidental feature of *W'* that it is a way things are not rather than the way things are. I have argued that the accidentalness of the property of being realized is extendible to the modal properties of possibly being realized and of not possibly being realized. Certainly it seems to be logically possible—not precluded by the principles of correct reasoning about modality—that a way-for-things-to-be that might have been realized might have been instead a way-for-things-to-be that could not have been realized, and that a way-for-things-to-be

that could not have been realized might have been instead a way-for-things-to-be that might have been realized. The friends of *B* modal logic commit themselves to the loaded claim that it is logically true that the property of possibly being realized (or of being a way things might have been) is an essential property of the actual world. The friends of *S4* modal logic commit themselves to the similarly loaded claim that it is logically true that the property of not possibly being realized is always an essential property of those worlds that have it. The friends of *S5* modal logic commit themselves to the double-barreled claim that it is logically true that both the properties of possibly being realized and of not possibly being realized are always essential properties of the worlds that have them. Yet all admit that the property of being realized is merely an accidental property that possible worlds can have or lack. What, then, is the rationale for their extremely strong versions of logical essentialism? Why should the modal properties of possibly being realized and of not possibly being realized be any less contingent or accidental, from the point of view of pure logic, than the non-modal properties of being realized and of not being realized? These alleged logical truths do not *seem* logically true. Indeed, the last two alleged logical truths, I have argued, are false. The first alleged logical truth, even if it is true, and even if it is necessarily true, does not seem logically true. Surely the burden of proof falls on the logical essentialists with respect to modal properties. We have just seen that the standard “proofs” of the characteristic *B*, *S4*, and *S5* axioms are in fact fallacious, since they assume that any possible world is essentially a possible world (or, in the case of *S4*, that anything that is not a possible world is essentially not a possible world). Whereas this may be trivially true in the generic sense of “possible world,” it simply begs the question in the modal sense. The reasoning involved in any purported justification of the contentious doctrine of logical essentialism with respect to modal properties cannot make use of such modal logics as *B*, *S4*, or *S5*—any more than induction can be justified to the Humean skeptic by citing inductive evidence. The systems *B*, *S4*, and *S5* for reasoning about what might have been are precisely what are at issue.

We friends of *T* modal logic are committed to the claim that it is logically true that the actual world has the property of possibly being realized, that as a matter of logic alone, the way things are is a way things might have been. Here we have something that is

transparently logically true. Quite plainly, anyone who cannot recognize the validity of an inference from an assertion that it must be that such-and-such to the assertion that such-and-such, does not know how to reason correctly about what must be; and anyone who cannot recognize the validity of an inference from an assertion that such-and-such to the assertion that it might have been that such-and-such does not know how to reason correctly about what might have been. Even the characteristic *B* principle, which may well be necessarily true, does not seem logically true. A proper justification for *B* as a system of modal logic, as opposed to a justification for *B* as a *metaphysical theory* of modality, would require not merely a defense of the truth of the essentialist doctrine that the actual world is necessarily possible, and not merely a philosophical argument that the doctrine is indeed a necessary truth, but a convincing case that the doctrine is, like the characteristic principle of *T*, required by logic and nothing more. Until such a justification is provided, modal reasoning in accordance with *B* is not to be recommended—except, of course, insofar as one is prepared to accept a commitment to a certain metaphysical theory. Even then, the *B* “axioms” would not be logical axioms, properly so-called, but metaphysical postulates or premises.

If the modal logical systems *B*, *S4*, and *S5* have never been satisfactorily justified, why are they almost universally accepted as correct systems for reasoning about what must be and what might have been? I have already cited several sources of the present confused state of affairs in contemporary philosophical logic. First, there is the generic-modal ambiguity in the phrase “possible world,” which has led to the widely accepted myths that the concepts of metaphysical necessity and possibility are defined in terms of, or constructed from, the concept of a possible world and that metaphysical modality is unrestricted modality. Equivocation between these two senses of “possible world” has led to the fallacious “proofs” of the characteristic *B*, *S4*, and *S5* principles. These fallacious arguments very likely owe something also to another source of confusion in contemporary philosophical logic: the widely adopted ostrich approach to modality, with its consequent misconstrual of “necessarily” as meaning actual necessity and “possibly” as meaning actual possibility. In fact, if the indexical sentential operator “actually” is added to the modal resources of a language, with appropriate logical axioms and restrictions governing its use in

modal reasoning, while retaining only the weak modal system  $T$  for the underlying logic of “necessarily” and “possibly,” exact analogues to the characteristic axioms and rules of  $T$ ,  $B$ ,  $S4$ , and  $S5$  emerge as trivial theorems for the special complex modal operators “actually necessarily” and “actually possibly.” In this sense,  $S5$  (as the logic of “actually necessarily” and “actually possibly”) is a subtheory of  $T$  plus the modal logic of “actually.”<sup>17</sup> Given its mis-construal of “necessarily” and “possibly,” the ostrich approach thus inevitably leads to the acceptance of  $S5$  as the correct logic for these modal operators.

My claim is this: the sort of consideration raised in Section I above *demonstrates* the invalidity of  $S4$  modal reasoning. I am not proposing a rejection of  $S4$  in an *ad hoc* manner, as merely an effective measure for avoiding the difficulty, with no further justification beyond the fact that it avoids the difficulty. The difficulty stems from a widely shared modal intuition, to the effect that some

---

<sup>17</sup>The observation made in the last two sentences derived in part from a fruitful discussion in Dubrovnik with Timothy Williamson (who does not fully endorse the views defended in this article). Williamson correctly observed that although infinitely iterated necessity and infinitely iterated possibility are modal operators for which the analogue of  $S4$  is derivable as a subtheory using only  $T$  as the underlying logic of “necessarily” and “possibly,” the analogues of  $B$  and  $S5$  are not thus derivable, since the infinitely iterated modalities replace ordinary accessibility by its ancestral, which is automatically transitive but which is not logically required to be symmetric if ordinary accessibility is not. Williamson wondered whether, on my view, there is any modal operator that is definable in terms of “necessarily,” and for which the analogue of  $S5$  is derivable as a subtheory using only  $T$  as the underlying logic of “necessarily.” The answer I proposed was: “actually necessarily.” (See also note 9 above.)

One characteristic axiom schema of the logic of “actually” is ‘actually  $\phi \supset \Box$  actually  $\phi$ ’. Another is ‘ $\phi \equiv$  actually  $\phi$ ’. Application of the rule of necessitation must be restricted to subsidiary proofs that do not invoke the latter axiom.

Williamson’s observation generates one serious difficulty for a claim that is often made in response to my arguments and which is closely bound to the myth that metaphysical modality is completely unrestricted: that the logic of necessity and possibility *has to be*  $S5$  because “necessarily” really means what I am calling “infinitely iterated necessity” and “possibly” really means what I am calling “infinitely iterated possibility.” The logic of what I am calling the “infinitely iterated modalities” would seem to be not  $S5$  but  $S4$ . (A more immediate difficulty with the suggested interpretation is its intrinsic implausibility. For example, it rejects the intuition that, necessarily, Woody might have originated from any wood that is only one molecule different from its original wood but could not have originated from entirely different wood, as not merely false but literally inconsistent.)

small variation in the origin of a material artifact is possible whereas complete variation is impossible. Even if one does not share this intuition, however, it should be quite obvious that the modal position of one (such as myself) who canonizes the intuition into metaphysical doctrine is at least coherent. The position cannot be summarily dismissed on logical grounds alone, as one would (rightly) dismiss the position of someone who proposes restricting the inference rule of *modus ponens* or denying the Law of Noncontradiction or rejecting the characteristic principles of *T*. If the modal position in question seemed not only false but incoherent, a proposal to reject *S4* modal logic solely on the basis of the modal intuition in question would indeed be drastic and poorly motivated. But the mere logical possibility, as opposed to the truth, of the modal intuition is beyond all reasonable doubt. Mere logical possibility, as opposed to truth, is what my argument against *S4* requires. The position outlined in Section I yields a *model* or *interpretation* that both respects the intended interpretation of the logical constants, including “necessarily” (see note 2), and invalidates *S4*. Due consideration of this difficulty makes it *intuitively* plain that *S4* modal reasoning involves a fallacy. Every attempt that I am aware of to retain *S4* modal logic in the face of this difficulty is distinctly counterintuitive.<sup>18</sup> The sort of consideration raised in Section I *exposes* a certain modal *fallacy*, that of inferring the iterated necessity claim “It must be that it must be that such-and-such” from the weaker claim “It must be that such-and-such.” Elsewhere I have called this “the fallacy of necessity iteration.” This fallacy is the very cornerstone of *S4* modal logic.

*University of California, Santa Barbara*

---

<sup>18</sup>By far the most popular such attempt is the proposal—made or suggested by Roderick Chisholm, Graeme Forbes, Anil Gupta, Saul Kripke, and Robert Stalnaker (to name but a few)—to replace standard modal semantics with some form or other of counterpart-theoretic modal semantics, as championed by David Lewis. (Kripke’s suggestion of a counterpart-theoretic treatment for philosophical problems of the sort engendered by Woody vis-à-vis the matter *m* is made more or less in passing, amid an emphatic rejection of counterpart theory for less problematic modal contexts. See note 12 above.) This alternative system of modal semantics allows for the retention of *S5* modal propositional logic, at a considerable cost. For an accounting of the costs involved, see “Modal Paradox: Parts and Counterparts, Points and Counterpoints.”

## SELECTED BIBLIOGRAPHY

- Adams, R. M., "Theories of Actuality," *Noûs* 8 (1974), pp. 211–231; also in Loux, pp. 190–209.
- Adams, R. M., "Primitive Thisness and Primitive Identity," *The Journal of Philosophy* 76 (1979), pp. 5–26.
- Adams, R. M., "Actualism and Thisness," *Synthese* 49 (1981), pp. 3–41.
- Bringsjord, S., "Are There Set Theoretic Possible Worlds?" *Analysis* 45 (1985), p. 64.
- Burke, M. B., "Essentialism and the Identity of Indiscernibles," *Philosophy Research Archives* 9 (1983), pp. 223–243.
- Carter, W. R., "Salmon on Artifact Origins and Lost Possibilities," *The Philosophical Review* 92 (1983), pp. 223–231.
- Chandler, H., "Plantinga and the Contingently Possible," *Analysis* 36 (1976), pp. 106–109.
- Chisholm, R., "Identity Through Possible Worlds: Some Questions," *Noûs* 1 (1967), pp. 1–8; also in Loux, pp. 80–87.
- Chisholm, R., "Parts as Essential to Their Wholes," *Review of Metaphysics* 26 (1973), pp. 581–603.
- Chisholm, R., "Mereological Essentialism: Some Further Considerations," *Review of Metaphysics* 28 (1975), pp. 477–484.
- Chisholm, R., *Person and Object* (La Salle, Ill.: Open Court, 1976).
- Forbes, G., "Origin and Identity," *Philosophical Studies* 37 (1980), pp. 353–362.
- Forbes, G., "On the Philosophical Basis of Essentialist Theories," *The Journal of Philosophical Logic* 10 (February 1981), pp. 73–99.
- Forbes, G., "Canonical Counterpart Theory," *Analysis* 42 (1982), pp. 33–37.
- Forbes, G., "Thisness and Vagueness," *Synthese* 54 (1983), pp. 235–259.
- Forbes, G., "Two Solutions to Chisholm's Paradox," *Philosophical Studies* 46 (1984), pp. 171–187.
- Forbes, G., *The Metaphysics of Modality* (Oxford, England: Oxford University Press, 1985).
- French, P., Uehling, T., and Wettstein, H., eds., *Midwest Studies in Philosophy XI: Studies in Essentialism* (Minneapolis, Minn.: University of Minnesota Press, 1986).
- Fumerton, R. A., "Chandler on the Contingently Possible," *Analysis* 38 (1978), pp. 39–40.
- Grim, P., "On Sets and Worlds: A Reply to Menzel," *Analysis* 46 (1987), pp. 186–191.
- Gupta, A., *The Logic of Common Nouns* (New Haven, Conn.: Yale University Press, 1980).
- Hazen, A., "Counterpart Theoretic Semantics for Modal Logic," *The Journal of Philosophy* 76 (1979), pp. 319–338.
- Kaplan, D., "Transworld Heir Lines," in Loux, pp. 88–109.
- Kaplan, D., "How to Russell a Frege-Church," *The Journal of Philosophy* 72 (1975), pp. 716–729; also in Loux, pp. 210–224.
- Kripke, S., "Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi," *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 9 (1963), pp. 67–96.
- Kripke, S., "Semantical Considerations on Modal Logic," *Acta Philosophica*

- Fennica* 16 (1963), pp. 83–94; also in L. Linsky, ed., *Reference and Modality* (New York, N.Y.: Oxford University Press, 1971), pp. 63–72.
- Kripke, S., “Identity and Necessity,” in M. Munitz, ed., *Identity and Individuation* (New York, N.Y.: New York University Press, 1971), pp. 135–164; also in S. Schwartz, ed., *Naming, Necessity, and Natural Kinds* (Ithaca, N.Y.: Cornell University Press, 1977), pp. 66–101.
- Kripke, S., *Naming and Necessity* (Cambridge, Mass.: Harvard University Press and Oxford, England: Basil Blackwell, 1972).
- Lewis, D., “Counterpart Theory and Quantified Modal Logic,” *The Journal of Philosophy* 65 (1968), pp. 113–126; also in Loux, pp. 110–128; also in Lewis, 1983, with substantive postscripts, pp. 26–46.
- Lewis, D., *Counterfactuals* (Cambridge, Mass.: Harvard University Press, 1973).
- Lewis, D., *Philosophical Papers I* (New York, N.Y.: Oxford University Press, 1983).
- Lewis, D., *On the Plurality of Worlds* (Oxford, England: Basil Blackwell, 1986).
- Loux, M., ed., *The Possible and the Actual* (Ithaca, N.Y.: Cornell University Press, 1979).
- Menzel, C., “On Set Theoretic Possible Worlds,” *Analysis* 46 (1986), pp. 68–72.
- McGinn, C., “On the Necessity of Origin,” *The Journal of Philosophy* (1976), pp. 127–135.
- McMichael, A., “A Problem for Actualism About Possible Worlds,” *The Philosophical Review* 92 (1983), pp. 49–66.
- Naylor, M. B., “A Note on David Lewis’s Realism About Possible Worlds,” *Analysis* 46 (1986), pp. 28–29.
- Noonan, H. W., “The Necessity of Origin,” *Mind* 92 (1983), pp. 1–20.
- Odegard, D., “On A Priori Contingency,” *Analysis* 36 (1976), pp. 201–203.
- Over, D., “The Consequences of Direct Reference,” *Philosophical Books* 25 (1984), pp. 1–7.
- Plantinga, A., “On Mereological Essentialism,” *Review of Metaphysics* 28 (1975), pp. 468–476.
- Plantinga, A., *The Nature of Necessity* (Oxford, England: Oxford University Press, 1974).
- Quine, W. V. O., “Worlds Away,” *The Journal of Philosophy* 73 (1976), pp. 859–863.
- Salmon, N., “How Not to Derive Essentialism from the Theory of Reference,” *The Journal of Philosophy* 76 (1979), pp. 703–725.
- Salmon, N., *Reference and Essence* (Princeton, N.J.: Princeton University Press and Oxford, England: Basil Blackwell, 1981).
- Salmon, N., “Fregean Theory and the Four Worlds Paradox: A Reply to David Over,” *Philosophical Books* 25 (1984), pp. 7–11.
- Salmon, N., “Impossible Worlds,” *Analysis* 44 (1984), pp. 114–117.
- Salmon, N., *Frege’s Puzzle* (Cambridge, Mass.: Bradford Books/The MIT Press, 1986).
- Salmon, N., “Reflexivity,” *The Notre Dame Journal of Formal Logic* 27 (1986), pp. 401–429.
- Salmon, N., “Modal Paradox: Parts and Counterparts, Points and Counterpoints,” in French, Uehling, and Wettstein, eds., 1986, pp. 75–120.
- Salmon, N., “Existence,” in J. Tomberlin, ed., *Philosophical Perspectives I: Metaphysics* (Atascadero, Calif.: Ridgeview, 1987), pp. 49–108.



- Salmon, N., "Review of David Lewis, *On the Plurality of Worlds*," *The Philosophical Review* 97 (1988), pp. 237–244.
- Sharvy, R., "Why a Class Can't Change its Members," *Noûs* 2 (1968), pp. 303–314.
- Sharvy, R., "Individuation, Essence and Plenitude," *Philosophical Studies* 44 (1983), pp. 61–70.
- Stalnaker, R., "Possible Worlds," *Noûs* 10 (1976), pp. 65–75; also in Loux, pp. 225–234.
- Stalnaker, R., "Counterparts and Identity," in French, Uehling, and Wettstein, eds., 1986, pp. 121–140.
- van Inwagen, P., "Two Concepts of Possible Worlds," in French, Uehling, and Wettstein, eds., 1986, pp. 185–213.
- Wilson, N. L., "Substances Without Substrata," *Review of Metaphysics* 12 (1959), pp. 521–539.

Princeton University

Department of Philosophy  
1879 Hall  
Princeton, New Jersey 08544 U.S.A.

February 3, 1987

Professor Nathan Salmon  
Department of Philosophy  
University of California, Santa Barbara  
Santa Barbara, CA 93106

Dear Nathan,

Enclosed are two papers, including the one in the Linsky volume too. The treatment of the relation  $R$  is on pp. 69-70 (section 2.1) of "Semantical Analysis" and on p. 64 of "Semantical Considerations". The same characterization of  $R$  is given in both places. As you see, there is no suggestion that  $S5$  is basic and the weaker systems come from some restricted conception.  $R$  is characterized in terms of truth and possibility of propositions in worlds. Notice also the discussion of the reduction axioms on p. 70 of "Semantical Analysis", and in particular of transitivity and  $S4$ .

One thing I do is, I now think, somewhat misleading. I should have stressed that the use of  $R$  does not make "possible" (as applied to worlds) into a two-place predicate, any more than, as you say, "is bald" is. Probably I only noticed this afterwards. Also, I should have stressed that strictly speaking, many of the worlds are not "possible" but only "possibly possible", and so on, unless we have  $S4$ .

By the time I gave the seminar I talked to you about I had definitely thought these points through, having seriously considered whether the conventional presupposition that the basic modal logic is  $S5$  is justified.

I am getting closer to thinking that your treatment of the ship is the correct solution. Certainly it is a very good piece of work. I am sorry if almost everyone is unable to see its virtues (you don't say quite that in the paper). As far as I can see, their counterarguments, as presented, are confused or circular. It was good talking to you. Talk to you about Russell, etc., some time.

Best,

  
Saul Kripke

Enc.