

# Which Modal Models are the Right Ones (for Logical Necessity)?

John P. BURGESS

ABSTRACT: Recently it has become almost the received wisdom in certain quarters that Kripke models are appropriate only for something like metaphysical modalities, and not for logical modalities. Here the line of thought leading to Kripke models, and reasons why they are no less appropriate for logical than for other modalities, are explained. It is also indicated where the fallacy in the argument leading to the contrary conclusion lies. The lessons learned are then applied to the question of the status of the formula  $\Diamond \exists x \exists y (x \neq y)$ .

Keywords: modality, models, logical necessity, modal logic, Kripke, Hintikka

## Introduction

If modern modal logic had picked up where mediaval modal logic left off, it would have begun with a well-developed distinction corresponding to what in current terminology would be called the distinction between *metaphysical* and *logical* necessity<sup>1</sup>. But as history actually unrolled, the technical side of modern modal was elaborated for many years before this elementary conceptual distinction became at all widely understood. Until the re-evaluation (at present only just begun) of the significance of technical results in the light of this crucial distinction is completed, existing work in modal logic can hardly contribute to the development of a good theory of logical necessity and therewith of logical consequence.

And even as this re-evaluation begins to be undertaken, there is a danger of its being muddled by what seems an increasingly widely held view that Kripke models, the main technical tool in almost all investigations in modal logic, are appropriate only for metaphysical modalities, and that a different notion, Kanger models, is required for logical modalities<sup>2</sup>. The present note reviews the reasons for rejecting this opinion<sup>3</sup>. §1 contains needed notational and terminological preliminaries. §2 is an exposition of a

---

<sup>1</sup> See the discussion of *consequentia* in Kneale and Kneale (1962), pp. 274ff, and especially of the distinction between *formal* and *material* consequence, the latter being, roughly speaking, enthymemes that become formally valid when a metaphysically necessary premise is added.

<sup>2</sup> The work of Kanger alluded to is his dissertation (Kanger (1957)). Until it was recently made available in his *Collected Works*, this item was very difficult to obtain. And the dissertation, once obtained, remains very difficult to read, mainly owing an accumulation of non-standard notation and terminology, even for notions like reflexive, symmetric, and transitive relations; but fortunately a readable account of Kanger's treatment of modality has recently been provided in Lindström (1998), where related ideas are found also in other early writers.

<sup>3</sup> In a recent paper (Burgess (1999)) I had occasion to mention this criticism of Kripke models in passing, quoting Jakko Hintikka there, as I will quote him here, as its most prominent and persistent advocate, and adding a couple of paragraphs of counter-criticism. Personal communications received since have convinced me that Hintikka's views are more widely shared than I had imagined, making a more detailed criticism of his criticisms desirable. I am especially indebted to Matthew McKeon, who is developing his own line of response, who brought to my attention in correspondence quite a number of published statements of views like Hintikka's. I have also profitted from correspondence with Sten Lindström.

line of thought leading to Kripke models, one of which there are traces in Kripke's own early publications<sup>4</sup>. I hope the exposition may be of some pedagogical value. §3 sketches an argument why, when necessity is understood as *truth by virtue of logical form*, Kripke's model theory should boil down to something simpler, and why it should *not* boil down to Kanger's model theory. §4 indicates how the considerations advanced in earlier sections bear on the vexed question of the status of the formula  $\diamond\exists x \exists y (x \neq y)$ <sup>5</sup>.

### 1. Use-mention confusions and how to avoid them

Before launching into the exposition, some preliminaries are needed. In all but the last section of this note, discussion will be confined to the level of sentential or propositional logic, or as I will call it (in an attempt to remain neutral), the logic of *statements*. Already at this level one frequently finds usages in the literature that invite a confusion of use and mention, which has plagued modern modal logic from its first beginnings.

Specifically, one often finds the letters  $p$ ,  $q$ ,  $r$ , and so on, employed in two quite different ways, exemplified by the following:

- (X) If  $p$  is any statement, then the negation of  $p$  is another statement, one which is true if  $p$  is false and false if  $p$  is true, and logically necessary if  $p$  is not logically possible, and logically possible if  $p$  is not logically necessary.
- (Y) The formula  $p \vee \sim p$  is *valid* because on *any* valuation (that is, any assignment of values  $\mathbf{T}$  and  $\perp$  to statement letters  $p, q, r \dots$ ) the formula receives the value  $\mathbf{T}$ . The formula  $p \ \& \ \sim p$  is *unsatisfiable* because there is *no* valuation on which the formula receives the value  $\mathbf{T}$ <sup>6</sup>.

What is the difference here? In (X),  $p$  is being *used* to refer to statements such as "Snow is white" or "Grass is green", and more specifically is being used as a variable ranging over such statements. In (Y),  $p$  and  $q$  and  $r$  are being *mentioned* as components out of which formulas are built up with the help of logical symbols such as  $\sim$  and  $\&$  and  $\vee$ . It would represent a confusion between use and mention, or if you will, between metalanguage and object language, to speak of formulas as true or false, or as logically necessary or logically possible — or inversely, to speak of statements as assigned value  $\mathbf{T}$  or assigned value  $\perp$ , or as valid or satisfiable.

In order to maintain the crucial distinction, it is not obligatory to use some special system of quotation and quasi-quotation marks. It is in fact difficult to devise a system of punctuation of manageable complexity that would mark all the distinctions there

---

<sup>4</sup> Especially Kripke (1959). To a limited extent, oral comments of Saul Kripke himself have seemed to me to confirm that his route to Kripke models was at least had some affinities with the route I will be discussing. But while I certainly wish to claim no special originality for this line of thought, neither would I wish to saddle Kripke with every opinion I may express in the course of expounding it.

<sup>5</sup> The question is especially pressed in Field (1989), pp. 116ff.

<sup>6</sup> My (X) and (Y) are not direct quotations from any specific writer. But one can find  $p$  and  $q$  used both as I will be using them, *and* as I will be using  $\alpha$  and  $\beta$ , and indeed also as I will be using  $A$  and  $B$ , throughout the crucial chapter VIII of Lewis and Langford (1959).

are to be drawn in this area; and the example of mediæval logicians, with their theories of *suppositio*, shows that the subtlest distinctions among different modes of employment of the same expression can be maintained without the use of any special signs at all. Nonetheless, some kind of notational distinction seems desirable as a precaution. My own solution will be to use a different style of letter,  $\alpha, \beta, \gamma$ , in contexts like (X), that is, for *statements*, and use the original,  $p, q, r$ , only in contexts like (Y), that is, for the *statement letters* out of which formulas are compounded. Besides, I will use a third style of letter,  $A, B, C$ , for the formulas compounded out of these sentence letters.

I will also make a related terminological distinction. The result of replacing the statement letters in a formula  $A$  with specific statements, such as “Snow is white” or “Snow is black” (with simultaneous replacement of logical symbols  $\sim, \&, \vee$ , and so on, by the logical operations of negation, conjunction, disjunction, and so on, that they are supposed to represent) I will call an *instantiation* of  $A$ . By contrast, I will call the result of replacing the statement letters in  $A$  by other formulas (whether sentence letters or compound formulas), a *substitution* in  $A$ . Thus an instantiation of  $A$  is the sort of thing that may be logically possible, or actually true, or logically necessary, but a substitution in  $A$  is the sort of thing that may be satisfiable, or assigned value  $\mathbf{T}$  by a valuation, or valid.

Now to ward off all possibility of confusion, there is a further point about statement letters that needs to be made. The standard aim of logicians at least from Russell onward has been to characterize the class all formulas *all of whose instantiations are true*. Thus, though Russell was a logical atomist, when he endorsed  $p \vee \sim p$  as law of logic, he did not mean to be committing himself only to the view that the disjunction of any *logically atomic* statement with its negation is true, but rather to be committing himself to the view that the disjunction of *any statement whatsoever* with its negation is true (or, to state once explicitly a qualification that will henceforth be tacitly understood, if not literally *any*, at least any within some class *closed under the logical operations for which one has introduced logical symbols*, and hence emphatically not limited to logical atoms). This has remained the standard employment of statement letters ever since, not only among Russell’s successors in the classical tradition, but also among the great majority of formal logicians who have thought classical logic to be in need of additions and/or amendments, including C. I. Lewis, the founder of modern modal logic.

With such an understanding of the role of statement letters, it is clear that if  $A$  is a law of logic, and  $B$  is any substitution in  $A$ , then  $B$  also is a law of logic. That is, if all instantiations of  $A$  are true, and  $B$  is a substitution in  $A$ , then all instantiations of  $B$  are true. This is simply because any instantiation of a substitution  $B$  is an instantiation of the original  $A$ . Thus it is that the rule of substitution applies not only in classical logic, but in standard, Lewis-style modal logics (as well as in intuitionistic, temporal, relevance, quantum, and other logics).

None of this is meant to deny that there may be circumstances where it is legitimate to adopt some other understanding of the role of statement letters. If one does so, however, it is indispensable to note the conceptual distinction, and highly advisable

to make a notational and terminological distinction<sup>7</sup>. Above all, it is important to recognize that one logician's proposed set  $\Gamma$  of formulas representing of laws of logic cannot be meaningfully compared with another logician's proposed set  $\Delta$  of formulas representing of laws of logic, if the two have different understandings of the roles of statement letters. If the first logician is adhering to some non-standard understanding of that role, and the second logician to the standard one, then it is not  $\Gamma$  itself but rather the set  $\Gamma^\#$  of those formulas *all substitutions in which* are in  $\Gamma$  that should be compared with  $\Delta$ <sup>8</sup>.

In the present study there is only one kind of non-standard employment of statement letters that will be of interest. Let us call the result of replacing the statement letters in a formula  $A$  by statements an *instantiation\** if *statement letters are replaced by statements that are logically atomic* (or, to state explicitly once a qualification that will henceforth be tacitly understood, if not literally logically *atomic*, at least without further logical structure *that can be represented only using whatever logical symbols one is using*), and *with distinct sentence are instantiated by statements that are logically independent*.. (Here  $n$  statements  $\alpha_1, \dots, \alpha_n$  are *independent* if all  $2^n$  combinations of truth values are possible.)

A formula fully indicates the logical form of its instantiations\* (insofar as it can be represented with the logical symbols one is using), but not of all its instantiations. For example, "Grass is green or snow is white" is an instantiation\* of  $p \vee q$ , while "Grass is green or grass is not green" is an instantiation of  $p \vee q$  that is not an instantiation\* thereof. It is, rather, an instantiation\* of  $p \vee \sim p$ , which is a substitution in  $p \vee q$ . There are contexts where it is of interest to consider, not the set of formulas all of whose instantiations are true, but rather the set of formulas all of whose instantiations\* are true<sup>9</sup>. This amounts to adopting a non-standard understanding of the role of statement letters, which I propose to indicate notationally by starring those letters. In the end I will use five styles of letters ( $\alpha, p, A, p^*, A^*$ ): for statements; for statement letters in their standard role; for formulas built up therefrom; for statement letters in a non-standard role; and for formulas built up therefrom. I fear that with any fewer styles of letter that this, there will be a danger of confusion.

## 2. The road to Kripke models

*Stage zero.* The idea in setting up a model theory is to introduce certain objects called *models*, and a certain relation between formulas and models called *holding* or *being valued*

<sup>7</sup> Such is done, for instance, by Arthur Prior, in his discussion of future contingents in (1967), chapter 7, where in addition to  $p, q, r$  for arbitrary statements he introduces  $a, b, c$ , for statements *not about the future*.

<sup>8</sup> This point is explained with admirable clarity by J. C. C. McKinsey, (1945), p. 85.

<sup>9</sup> One such context would be in an attempt at a reconstruction of logical atomist thought. Such a reconstruction has been attempted by Nino Cocchiarella, e.g. in (1984). While critics of Kripke models sometimes cite Cocchiarella as if his work supported their position, there is this difference, that the critics *never acknowledge at all that they are working with a non-standard understanding of the role of statement letters*. My counter-criticism will be directed against this lack of acknowledgment.

$\top$  (as opposed to *failing* or *being valued*  $\perp$ ), and to do this in such a way that every formula that holds in all models will have all its instantiations true (*soundness*)<sup>10</sup> and every formula all of whose instantiations are true will hold in all models (*completeness*). Thus in characterizing which formulas are laws of logic, the indefinitely large totality of all instantiations is replaced by the totality of all models, which may be more tractable, at least theoretically, if not practically.

In seeking a suitable model theory for modal logic, we begin with classical statement logic, with no modalities. At this level, since whether an instantiation of a formula  $A$  is true or not depends only on whether the statements  $\alpha_1, \dots, \alpha_n$  that instantiate the statement letters  $p_1, \dots, p_n$  in  $A$  are true or not, we may simply take as our models all *valuations* or assignments of values  $\top$  or  $\perp$  to statement letters, extending any such valuation  $V$  to complex formulas according to the usual rules:

- |      |               |          |                |                                 |
|------|---------------|----------|----------------|---------------------------------|
| (1a) | $V(\sim A)$   | = $\top$ | if and only if | not $V(A) = \top$               |
| (2a) | $V(A \& B)$   | = $\top$ | if and only if | $V(A) = \top$ and $V(B) = \top$ |
| (3a) | $V(A \vee B)$ | = $\top$ | if and only if | $V(A) = \top$ or $V(B) = \top$  |

The indefinitely large totality of all instantiations is adequately represented (in the sense that soundness and completeness hold) by the set of the  $2^n$  possible valuations of  $n$  letters.

*Stage one.* We next allow modalities, but no nesting of modalities, or modalities inside modalities. At this level, whether an instantiation of a formula  $A$  is true or not depends not only on which combination of truth values the statements  $\alpha_1, \dots, \alpha_n$  that instantiate the statement letters  $p_1, \dots, p_n$  in  $A$  actually have, but also on which of the  $2^n$  combinations of truth values are *possible*.

Consider, for example,

$$A = p_1 \& p_2 \& \diamond(p_1 \& \sim p_2) \& \diamond(\sim p_1 \& p_2)$$

If (i) we take as  $\alpha_1$  “Grass is green” and as  $\alpha_2$  “Snow is white”, then the instantiation of  $A$  we obtain is true, but not so if (ii) we take as  $\alpha_1$  “Grass is green” and as  $\alpha_2$  “Grass is green or snow is white”, or if (iii) we take as  $\alpha_1$  “Grass is green” and as  $\alpha_2$  “Grass is green and snow is white”. In all three cases (i)-(iii) the first two conjuncts will be true, but the third will be false in case (ii) and the fourth in case (iii). What we need to take for a model at this stage is, therefore, a valuation  $V_0$  representing what, in a given instantiation, the actual combination of truth values of the relevant statements would be, *together with* a set  $R_0$  of valuations representing the “realm of possibility”, that is, the *possible* combinations of truth values of those statements.

---

<sup>10</sup> Specifically, this version of soundness corresponds to the first implication in formula (2) in Mario Gomez-Torrente’s (2003) introduction to this symposium: the requirement that every argument that is “valid” in the technical sense of preserving truth in all structures, should be genuinely logically valid.

Already perhaps at this stage (and certainly by the next) the notion of model is complicated enough that one may wish to draw a picture: a dot representing the actual valuation  $V_0$  and oval containing many dots, representing the set  $R_0$  and the valuations belonging to it. Such a picture, once drawn, suggests a modification of the notion of model, which proves useful when we go on to later stages. With this modification, a model consists of a *point* or *index*  $x_0$ , representing actuality, to which is assigned a valuation  $V_0 = \mathcal{V}(x_0)$  and also a set  $R_0 = \mathcal{R}(x_0)$  whose elements are further points or indices  $x$  to each of which a valuation  $\mathcal{V}(x)$  is attached. A formula  $A$  without modalities will be said to hold or fail, or to have the value  $\top$  or  $\perp$ , at a point  $x$  if and only if it does so at the valuation  $\mathcal{V}(x)$  attached to  $x$ . Writing  $x \models A$  to indicate  $A$  holds at  $x$ , we thus have

- |      |                            |                |                                 |
|------|----------------------------|----------------|---------------------------------|
| (0b) | $x \models p$              | if and only if | $(\mathcal{V}(x))(p) = \top$    |
| (1b) | $x \models \sim A$         | if and only if | not $x \models A$               |
| (2b) | $x \models (A \ \& \ B)$   | if and only if | $x \models A$ and $x \models B$ |
| (3b) | $x \models (A \ \vee \ B)$ | if and only if | $x \models A$ or $x \models B$  |

And when  $x$  is distinguished index  $x_0$ , representing actuality, we have

- |      |                        |                |   |
|------|------------------------|----------------|---|
| (4b) | $x \models \Box A$     | if and only if | $y \models A$ for all $y \in \mathcal{R}(x)$  |
| (5b) | $x \models \Diamond A$ | if and only if | $y \models A$ for some $y \in \mathcal{R}(x)$ |

for any formula  $A$  without modalities. A formula counts as holding or failing in the model as a whole according as it holds or fails at the distinguished index  $x_0$ .

*Stage two.* If we allow nesting of modalities two deep, so that we have modalities inside modalities, but no modalities inside modalities inside modalities, then the notion of model must be further elaborated. Besides the index  $x_0$  representing actuality, with its attached valuation  $\mathcal{V}(x_0)$  representing actual truth and its attached set  $\mathcal{R}(x_0)$  representing the realm of actual possibility, we will need for each  $x$  in  $\mathcal{R}(x_0)$ , representing an actual possibility, not only a valuation  $\mathcal{V}(x)$  representing truth relative to  $x$  but now also a set  $\mathcal{R}(x)$  representing the realm of possibility relative to  $x$ . And for any  $y$  in any  $\mathcal{R}(x)$ , representing a possible possibility, we will need a valuation  $\mathcal{V}(y)$ . The definition of holding and failing is unchanged, except that clauses (4b) and (5b) now apply to the  $x$  in  $\mathcal{R}(x_0)$  as well as  $x_0$  provided  $A$  contains no modalities, and for  $x = x_0$  apply to  $A$  having modalities but no modalities inside modalities, as well as to  $A$  containing no modalities.

*Stage three.* It should be clear how the pattern continues as we allow deeper and deeper nesting of modalities. If we wish to allow unlimited nesting, we will in the end need the following apparatus: a set  $X$  indices, a distinguished index  $x_0$ , an assignment of a set  $\mathcal{R}(x)$  of indices to each index  $x$ , and an assignment of a valuation  $\mathcal{V}(x)$  to each index. Clauses (0b)-(5b) above will then define holding or failing at an index in a

model, holding or failing the model as a whole being holding or failing at its distinguished index.

The definition may be restated in terms of the relation  $Rxy$  which an index  $x$  bears to all and only those indices  $y$  in  $\mathcal{R}(x)$ . In these terms, a model consists of a set  $X$  of indices, a distinguished index  $x_0$ , a two-place relation  $R$  on indices, and an assignment of a valuation  $\mathcal{V}(x)$  to each index. Thus the final clauses in the definition above would read

$$\begin{array}{ll} (4b) & x \models \Box P \qquad \text{if and only if} \qquad y \models P \text{ for all } y \text{ with } Rxy \\ (5b) & x \models \Diamond P \qquad \text{if and only if} \qquad y \models P \text{ for some } y \text{ with } Rxy \end{array}$$

One can similarly replace the assignment of valuations to indices by the a two-place function  $V(x, p) = (\mathcal{V}(x))(p)$  giving the value  $\top$  or  $\perp$  of a statement letter  $p$  at an index  $x$ .

It may be noted that we have let go of the idea that every index  $x$  should either be the distinguished index  $x_0$ , or be  $R$ -related to  $x_0$ , or be  $R$ -related to something  $R$ -related to  $x_0$ , and so on. In terms of the heuristic line of thought that has led us this far, we are allowing extraneous indices that correspond neither to actuality, nor to actual possibility, nor to actually possible possibility, and so on. But a little thought shows that the presence of extraneous indices makes no difference to whether a formula holds in a model or not. Since the goal in setting up a model theory is not that every piece of apparatus in the model should correspond to something in a given heuristic line of thought, but only that holding in all models should in the end coincide with having all instantiations true, extraneous indices may be allowed, if that makes the definition of model easier to state, as it does.

Thus we have arrived at Kripke models in the form in which they are usually presented in the literature<sup>11</sup>, apart from the picturesque but potentially misleading terminology (subsequently regretted by Kripke himself) that calls indices possible *worlds*. Note that along the way we have made virtually *no* assumptions about the kind of necessity involved. The apparatus is intended to be flexible, and applicable to many kinds of necessity, distinctions among different kinds being reflected not in differences in the basic set-up  $\langle X, x_0, R, \mathcal{V} \rangle$ , but rather in special conditions that may be imposed on the relation  $R$ . For instance, with *alethic* modalities, that is, with notions of necessity that imply truth, as do both metaphysical and logical necessity, the actual state of things must count among the possible states of things, so one will want to impose the condition that each index be in its own “realm of possibility”, that is,  $x \in \mathcal{R}(x)$  or  $Rxx$ , making the relation  $R$  *reflexive*.

### 3. Hintikka's fallacy

Of course, when the model theory was originally introduced, the first task was to correlate various conditions on  $R$  with various axioms of modal logic that had been con-

---

<sup>11</sup> See for instance the standard reference Bull and Segerberg (1984).

sidered in the literature from C. I. Lewis onwards. Of special interest here will be the axioms system **S5**, which implies that every formula with nested modalities is equivalent to one without. As is well known, this system turns out to correspond to imposing the conditions of reflexivity, symmetry, and transitivity on  $R$ , making it an equivalence relation. If extraneous indices are eliminated, as they always may be, we are left with models in which *all* indices are  $R$ -related, so that any mention of  $R$  may be dropped, and we have

$$(4c) \quad x \models \Box A \quad \text{if and only if} \quad y \models A \text{ for all } y$$

$$(5c) \quad x \models \Diamond A \quad \text{if and only if} \quad y \models A \text{ for some } y$$

In fact, the indices themselves are dispensable and we may go back to working directly with valuations. That is, we get equivalent results if we simply take as a model a valuation  $V_0$  together with a set  $R_0$  of valuations, with  $V_0 \in R_0$  and we get

$$(4d) \quad V(\Box A) = \mathbf{T} \quad \text{if and only if} \quad W(A) = \mathbf{T} \text{ for all } W \text{ in } R_0$$

$$(5d) \quad V(\Diamond A) = \mathbf{T} \quad \text{if and only if} \quad W(A) = \mathbf{T} \text{ for some } W \text{ in } R_0$$

These clauses make the value of  $\Box A$  or  $\Diamond A$  independent of  $V$ , since the righthand side of the definition does not mention  $V$ .

What may in the end be the correct logic of metaphysical modalities is presumably a question for metaphysicians, but the determination of the correct logic of logical modalities is a question for logicians alone. A good case can be made that, at least when logical necessity is understood as *truth by virtue of logical form alone* (as contrasted with, say, *verifiability by virtue of logical form alone*), **S5** is the right system, and the simple model theory for it that I have just indicated the right model theory<sup>12</sup>.

The case may be conveniently put by using starred statement letters. One can then argue as follows. Let us consider a formula  $\Box A^*$ . Its instantiations\* will be statements

it is logically necessary that  $\alpha$

where where  $\alpha$  is an instantiation\* of  $A^*$ . Now we are understanding the displayed condition to mean

any statement  $\beta$  of the same logical form as  $\alpha$  is true

---

<sup>12</sup> In his introduction, Mario Gomez-Torrente says that the abstract Tarskian method of defining a notion of structure, truth in a structure, and consequence as preservation of truth in all structures, can be applied to languages with extend those of ordinary logic. The present paper deals, mainly at the level of the logic of statements, with the problem of what is the best notion of structure when the extension adds 'it is logically necessary that', and assumes it to be a logical operator, and hence an operator to be represented when representing the logical form of statements involving it. The argument to follow depends crucially on this assumption.



So since a statement  $\beta$  is of the same logical form as an instantiation\*  $\alpha$  of a formula  $A^*$  if and only if  $\beta$  is also an instantiation\* of  $A^*$ , it follows that an instantiation\* of  $\Box A^*$  will be true if and only if all instantiations\* of  $A$  are true. Similarly, an instantiation\* of  $\Diamond A^*$  will be true if and only if some instantiation\* of  $A$  is true. Since we want the totality of models to represent the totality of instantiations\*, we will want to have

$$(4e^*) \quad V(\Box A^*) = \mathbf{T} \quad \text{if and only if} \quad W(A^*) = \mathbf{T} \text{ for all valuations } W$$

$$(5e^*) \quad V(\Diamond A^*) = \mathbf{T} \quad \text{if and only if} \quad W(A^*) = \mathbf{T} \text{ for some valuation } W$$

(All valuations are relevant because the notion of instantiation\* requires that the statements replacing distinct statement letters be independent.) Thus we may take a model to consist of a valuation  $V_0$ , representing actuality, and the set  $R_0$  of *all* valuations — or indeed, leaving the latter to be understood, we may take a model to consist simply of a valuation  $V_0$ , the same as for classical statement logic without modalities.

This, of course, does not complete the analysis if we are interested in comparing our results with the systems of C. I. Lewis and his followers, who adopted the *standard* understanding of the role of statement letters. As we said early on, for meaningful comparison to be possible, what we must then consider is not the set  $\Gamma$  of formulas  $A$  that are valid according to definitions (4d), (5d), but rather the set  $\Gamma^\#$  of formulas  $P$  *all of whose substitutions* are valid according to those definitions. And so we must inquire what is the effect of substituting formulas  $A_1^*, \dots, A_n^*$  for the statement letters  $p_1, \dots, p_n$  in a given formula  $A$ .

Evidently, a given valuation  $V$  of the statement letters  $p_1^*, \dots, p_n^*$  will give rise to a valuation  $V^\dagger(p_i) = V(A_i^*)$ . Thus a model consisting of a valuation  $V_0$  together with the set of all valuations will give rise to a model consisting of the valuation  $V_0^\dagger$  together with the set  $R_0$  of all valuations of form  $W^\dagger$ , which in general will be a proper subset of the set of all valuations. (For instance, if  $A_1^*$  is  $p_1^* \ \& \ \sim p_1^*$ , no valuation  $W$  will have  $W^\dagger(p_1) = \mathbf{T}$ .) Since we want to consider *all* substitutions, we must consider *all* models obtainable in this way. And thus we are led to the notion of a model consisting of a designated valuation and a set of valuations, with the clauses (4d) and (5d) above. The point to notice here is that we were forced to move from considering *all* valuations as in (4e\*), (5e\*) to considering only all valuations *in some designated set*  $R_0$  by a desire *to restore the standard understanding of the role of statement letters*.

At a purely formal level, one could be led to the same move in a quite different way, if one wished to *retain* the non-standard understanding of the role of statement letters, but wished to *replace* the notion of logical necessity by some weaker notion, such as metaphysical necessity. The correlative notion of possibility would then be stronger, so that while every valuation represents a logical possibility, it may be that not every valuation represents a metaphysical possibility, but only those in some distinguished set  $R_0$ . Of course, if one maintains, as I have urged, a notational distinction

between standard statement letters  $p$  and formulas  $\mathcal{A}$  built up from them on the one hand, and non-standard statement letters  $p^*$  and formulas  $\mathcal{A}^*$  built up from them on the other hand, what one would be led to would be, not literally (4d) and (5d) above, but rather the following:

$$\begin{array}{llll} (4d^*) & \mathcal{V}(\Box \mathcal{A}^*) & = \mathbf{T} & \text{if and only if } \mathcal{W}(\mathcal{A}^*) = \mathbf{T} \text{ for all } \mathcal{W} \text{ in } R_0 \\ (5d^*) & \mathcal{V}(\Diamond \mathcal{A}^*) & = \mathbf{T} & \text{if and only if } \mathcal{W}(\mathcal{A}^*) = \mathbf{T} \text{ for some } \mathcal{W} \text{ in } R_0 \end{array}$$

Similarly, with the notational distinctions I have been urging, we saw above that a case could be made, with logical modalities, for (4e\*) and (5e\*), based on the fact that two special instantiations of the same formula have the same logical form. But no such case can be made for

$$\begin{array}{llll} (4e) & \mathcal{V}(\Box P) & = \mathbf{T} & \text{if and only if } \mathcal{W}(P) = \mathbf{T} \text{ for all valuations } \mathcal{W} \\ (5e) & \mathcal{V}(\Diamond P) & = \mathbf{T} & \text{if and only if } \mathcal{W}(P) = \mathbf{T} \text{ for some valuations } \mathcal{W} \end{array}$$

since two instantiations of the same formula may have very different logical forms. (To cite the extreme example, *all* statements are instantiations of  $p$ , but not all statements have the same logical form.) And indeed (4e) and (5e) are flatly incompatible with a standard understanding of the role of statement letters, since they lead to failures in the rule of substitution. (For  $\Diamond p$  is valid, or valued  $\mathbf{T}$  in all valuations, while  $\Diamond(p \ \& \ \sim p)$  is not.)

Of course, what is crucial is not that one distinguish *notationally* between (4d) and (4d\*), or between (4e) and (4e\*), let alone that one adopt the particular starring convention I have used here. What is crucial is that one distinguish *conceptually* between statement letters thought of as representing *arbitrary* statements, and statement letters thought of as representing *independent atomic statements*. With the former, standard understanding, the restriction to a *subset* of all valuations in (4d) and (5d) need have nothing to do with a switch from logical to any kind of non-logical modalities, since it is required even for logical modalities, simply a reflection of the fact that with logically complex, logically interrelated statements instantiating the statement letters, not all combinations of truth values may be logically possible. And with the standard understanding, (4e) and (5e) represent a sheer blunder. By contrast, (4e) and (5e) with the non-standard understanding of the role of statement letters — which I would write as (4e\*) and (5e\*) — is appropriate for logical modalities; whereas (4d) with the same understanding — which I would write (4d\*) and (5d\*) — will be appropriate for non-logical modalities, such as metaphysical modalities.

This brings us to Kanger. His proposal for a model theory for logical modalities amounts, at the level of statement logic, to extending the usual definition from classical logic by adding clauses (4e) and (5e). He also says, in effect, that if one wanted to consider some weaker notion of necessity, the appropriate clauses would be (4d) and (5d), or something more elaborate involving a relation  $R$ , pointing in the direction of Kripke models.

Hintikka, writing after the publication of Kripke's work, goes further, and explicitly criticizes Kripke models as inappropriate for logical modalities precisely insofar as they differ from Kanger models for logical modalities by using clauses like (4d) and (5d) rather than like (4e) and (5e). On the strength of a very superficial analogy with the model theory of second-order logic, he calls Kripke models based on (4d) and (5d) "non-standard" and Kanger models based on (4e) and (4f) "standard". The end result is that we find Hintikka writing as follows:

...in its usual form, the so-called Kripke semantics is not the correct semantics for logical modalities either. As has been pointed out repeatedly, Kripke semantics, unlike e.g. the variant possible-worlds treatments by Kanger, is analogous to the nonstandard interpretations of higher-order logics, which is not equivalent to the intended standard interpretation of these logics. In other words, the so-called Kripke's semantics does not provide us with the right model theory of the logical (conceptual) necessities...<sup>13</sup>

Here Hintikka wholly fails to note that Kripke was working, like virtually every other modal logician beginning with C. I. Lewis, with a standard conception of the role of statement letters, for which Kanger's model theory, with its failure of the rule of substitution, is wholly inappropriate regardless of the kind of modality involved. His criticism of Kripke models is thus fallacious. Hence the title of this section<sup>14</sup>.

#### 4. *The size of the universe*

So much for modal statement logic. Turning to *quantified* modal logic, the first problem one faces in attempting to determine which is the right axiomatic system or the right model theory for logical modalities is the difficulty of making sense of quantification into context of logical, as contrasted with metaphysical, modality. To make sense of this notion one would have to make sense of the notion of an open statement, such as "x is green" being logically necessary *of a thing*, regardless of how or whether that thing is named or described in language<sup>15</sup>.

This problem, however, seems to pertain only to *de re* modality, with modalities inside quantifiers, and not *de dicto* modality, with quantifiers inside modalities. Hence one might at least hope to settle questions about the latter, more restricted topic. Even here, however, there is trouble, at least if identity is admitted. For there has been a controversy over the status of the formula

$$(1) \Diamond \exists x \exists y (x \neq y)$$

---

<sup>13</sup> Hintikka and Sandu (1995), p.281. As the word "repeatedly" in the quoted passage suggests, similar passages can be found in earlier works of Hintikka's, going all the way back to (1980).

<sup>14</sup> It should be emphasized that the fallacy does not consist simply in being interested in an approach to modality that implicitly involves a non-standard conception of the role of variables, as in Kanger (1957), but in confusing that approach with more standard approaches, and especially in criticizing work based on more standard approaches without recognizing the difference in conception, as in Hintikka (1980) and Hintikka and Sandu (1995).

<sup>15</sup> This problem was raised by Quine in response to the very first, purely formal, experiments in combining quantifiers and modalities. It is argued in Burgess (1998) that Quine's criticisms have never been adequately answered, so far as they pertain specifically to *logical* modalities.

Does this formula represent a law of logic, or not? The question is subtle enough that disagreement might be expected in any case, but discussion of the question has unfortunately been further muddled by Hintikka's and a kindred fallacy, which suggest two short, simplistic arguments for an affirmative answer to the question.

The first argument begins with the assumption that if one is interested in *logical* modalities, then at the level of statement logic, the correct notion of modal model will be the totality of all classical models for statement logic (that is, all valuations), with one of these distinguished as representing actuality, and reasons by analogy that at the level of quantification theory the correct notion of modal model will again be the totality of classical models, this time for quantification theory, again with one of these distinguished. Since the totality of classical models will include ones where

$$(2) \quad \exists x \exists y (x \neq y)$$

holds, (1) will hold in any modal model, or in other words, will be valid.

Clearly the initial assumption in this argument is simply Hintikka's fallacy, and so the argument is to be rejected along with the latter.

The second argument is that even if there is some finite upper bound on how many things there are, pure *logic* cannot teach us what this bound is, so it remains *logically* possible that there are more things than any given bound, and in particular more than two. Moreover, the foregoing considerations are themselves logical in character, so it is a law of logic that it is logically possible that there are more than two things.

The fallacy here resembles Hintikka's fallacy in that it involves a confusion over the role of letters appearing in formulas, only in this case it is not statement letters  $p, q, r$ , but rather individual variables  $x, y, z$  whose role is at issue. That it is logically possible for there to be more than two things may be in some sense a law of logic, but it is *not* what (1) represents on a standard understanding of the role of individual variables, since on such an understanding they do not range over the totality of all things, but rather over some non-empty domain or universe of discourse.

A third argument now suggests itself. Let us write  $x^+, y^+, z^+$ , and so on, for variables understood as ranging over *all* things. Then in this notation (1) in effect amounts to

$$(3) \quad \exists x^+ P x^+ \rightarrow \Diamond \exists x^+ \exists y^+ (P x^+ \ \& \ P y^+ \ \& \ x^+ \neq y^+)$$

Here the one-place predicate  $P$  represents belonging to the domain, and the antecedent of (3) corresponds to the standard assumption that the domain is non-empty. Now the argument would run that, even *without* this antecedent, the consequent must hold, since

$$(4) \quad \exists x^+ \exists y^+ (P x^+ \ \& \ P y^+ \ \& \ x^+ \neq y^+)$$

is satisfiable.

This last argument is just Hintikka's fallacy again. With a standard understanding of the role of predicate letters, (3) cannot represent a law of logic unless every substitution does, while clearly the substitution resulting from replacing  $Px^+$  by

$$Qx^+ \ \& \ \sim \exists \tilde{x}^+ (Q\tilde{x}^+ \ \& \ \tilde{x}^+ \neq x^+)$$

does not.

Whether (1) will represent a law of logic on the standard understanding of the role of individual variables depends not on whether it is logically possible for there to exist more than one element in the domain of all things, nor yet on whether for any domain specified by a logically atomic open statement it is logically possible for there to exist more than one element in the domain thus specified, but rather on whether it is the case that for *any* domain it is logically possible that there should be more than two elements in *it*.

But what does *that* mean? The difficulties in making sense of such a condition are familiar ones (or the second-order analogues of what at the first-order level are familiar ones). Thus there is no internal logical self-contradiction in the statement

(3) The set of female twentieth-century British prime ministers has more than one element.

For it is political history, not pure logic, that teaches us that (3) is false. But by contrast, the statement

(4) The set whose one and only element is Margaret Thatcher has more than one element.

is a logical absurdity. What, then, are we to say of the domain which is at once the set of female twentieth-century British prime ministers and the set whose one and only element is Margaret Thatcher? Is "X has more than one element" logically possible of *it*, or not?

My aim here is not to argue that the correct answer to this question is negative, any more than to argue that the correct answer is affirmative; nor yet is my aim to argue that the question is unanswerable. My aim is only to argue that the question of the status of (1) cannot be answered by the kind of short, simplistic arguments I have been considering, and that any progress on the question will require careful attention to the issue of what one conceives the role of various letters and symbols in formulas to be. Without close attention to such issues, the advance of modal logic on this and other fronts threatens to be derailed by Hintikka's fallacy and its kin. Confusions of metalanguage and object language, of use and mention, against which Quine so early and so often warned, remain a danger even today, alas.

## BIBLIOGRAPHY

- Bull, R. A. and K. Segerberg (1984), "Basic Modal Logic," in D. Gabbay and F. Guenther (eds.), *Handbook of Philosophical Logic*, vol. 2: *Extensions of Classical Logic*, Dordrecht: Reidel, 1-88.
- Burgess, J. P. (1998), "*Quinus ab Omne Nervo Vindicatus*," in *Meaning and Reference: Canadian Journal of Philosophy Supplement*, vol. 23, edited by A. A. Kazmi, 25-65.
- (1999), "Which Modal Logic is the Right One?," *Notre Dame Journal of Formal Logic*, vol. 40, 81-93.
- Cocchiarella, N. (1984), "Philosophical Perspectives on Quantification in Tense and Modal Logic," in D. Gabbay and F. Guenther (eds.), *Handbook of Philosophical Logic*, vol. 2: *Extensions of Classical Logic*, Dordrecht: Reidel, 309-54.
- Field, H. (1989), *Realism, Mathematics and Modality*, Oxford: Basil Blackwell.
- Gomez-Torrente, M. (2003), "Logical Consequence and Logical Expressions", *Theoria*, this volume.
- Hintikka, J. (1980), "Standard vs. Nonstandard Logic: Higher-Order, Modal, and First-Order Logics," in E. Agazzi (ed.), *Modern Logic — A Survey*, Dordrecht: Reidel, 183-96.
- Hintikka, J. and G. Sandu (1995), "The Fallacies of the New Theory of Reference," *Synthese*, vol.104, 245-83.
- Kanger, S. (1957), *Provability in Logic*; Acta Universitatis Stockholmensis, Almqvist and Wiksell, Stockholm; also in Ghita Holmström-Hintikka, Sten Lindström, and Rysiek Sliwinski (eds.), *Collected Papers of Stig Kanger with Essays on his Life and Work*, volume 1, Synthese Library volume 303, Dordrecht: Kluwer, 2001.
- Kneale, W. and M. Kneale (1962), *The Development of Logic*, Oxford: Clarendon Press.
- Kripke, S. (1959), "A Completeness Theorem in Modal Logic," *The Journal of Symbolic Logic*, vol. 24, 1-24.
- Lewis, C. I. and C. H. Langford (1959), *Symbolic Logic*, (2nd ed.) New York: Dover,.
- Lindström, S. (1998), "Kanger's Early Semantics for Modal Logic," in P. W. Humphreys and J. H. Fetzer (eds.), *The New Theory of Reference: Marcus, Kripke, and Its Origins*, Dordrecht: Kluwer, 203-33.
- McKinsey, J. C. C. (1945), "On the Syntactical Construction of Systems of Modal Logic," *The Journal of Symbolic Logic*, vol. 10, 83-94.
- Prior, A. N. (1967), *Past, Present and Future*, Oxford: Clarendon Press.

**John P. BURGESS**, Ph.D. in logic, Berkeley, and professor and director of undergraduate studies in philosophy, Princeton, is the author of numerous articles on logic and foundations of mathematics, and (with Gideon Rosen) of *A Subject With No Object* (Oxford, 1997) His other philosophical interests include metaethics and 'pataphysics.

**Address:** Department of Philosophy, Princeton University, Princeton, NJ 08544-1006, USA. E-mail: [jburgess@princeton.edu](mailto:jburgess@princeton.edu)