Quantified Modality and Essentialism

Saul A. Kripke

This is the accepted version of the following article:

Kripke, S. A. (2017), Quantified Modality and Essentialism. Noûs, 51: 221-234. doi: 10.1111/nous.12126,

Which has been published in final form at

https://doi.org/10.1111/nous.12126.

Quantified Modality and Essentialism[1]

Saul A. Kripke

It is well know that the most thoroughgoing critique of modal logic has been that of W.V. Quine. Quine's position, though uniformly critical of quantified modal systems, has nevertheless varied through the years from extreme and flat rejection of modality to a more nearly moderate critique.

At times Quine urged that, for purely logico-semantical reasons based on the problem of interpreting quantified modalities, a quantified modal logic is impossible; or, alternatively, that it is possible only on the basis of a queer ontology which repudiates the individuals of the concrete world in favor of their ethereal intensions. Quine has also urged that even if these objections have been answered, modal logic would clearly commit itself to the metaphysical jungle of Aristotelian essentialism; and this essentialism is held to be a notion far more obscure than the notion of analyticity upon which modal logic was originally to be based.

---

[1] There is a long story behind this paper. It was written for a seminar given by Quine himself in the academic year 1961/2, and discussed in class over a period of several weeks. Some years later, I was surprised to hear from Héctor-Neri Castañeda that it had received wider circulation among philosophers. I think I didn't even have a copy of the paper at the time, having handed my own to Quine for grading. Castañeda suggested I submit the paper to *Noûs* and located the original version – which contains some markings in the margins, presumably due to Quine – at the Harvard Philosophy Department library. I agreed to submit it to *Noûs*, though I mentioned in a letter that it needed some corrections. The paper was accepted for publication in 1966 and the idea was to include a reply by Quine as well, but this never came about and the final version was not submitted (until now). Recently, some people at the Saul Kripke Center, and also Graham Priest and Philip Bricker, read it and thought it should still be published. So I resubmitted it to *Noûs* and was told by the editor that it would be accepted if it hadn't already been accepted!

I would put things somewhat differently now, though the technical result is, as I say in the paper, impeccable. Some footnotes will reflect this (all footnotes, except footnote 18, are new). Unfortunately, I don't have a copy of the letter mentioning the corrections I thought were needed then. But by 1966 I had developed many of the philosophical views in *Naming and Necessity*, and would not think of this as the primary way of interpreting essential predication. Maybe that is what I had in mind.

If I have understood correctly the revised version of Quine's *From a Logical Point of View* (henceforth, *LPV*),[2] only the last objection, the objection to modal "essentialism," is retained in this most recent statement. But the objectionable features of essentialism still seem, in Quine's opinion, to be sufficient reason to declare, "so much the worse for quantified modal logic."[3] (*LPV*: 156)

The present paper attempts to analyze this objection; and my conclusion, indeed, is that Quine's fears of the notion, though by no means groundless, nevertheless do not count as any objection to quantified modal logic, but only to certain applications thereof. I shall state these points with greater precision after a brief review of some fundamental concepts.

Let us initially make a distinction between the *pure* and the *applied* semantics of a formal system of modal logic, say the quantified modal logic that I have called S5. In Kripke (1959a) I have given a semantical treatment of this system and have shown it to be complete; in fact, the entire argument can be carried out within Zermelo set theory. But the semantics in that paper was not intended *merely* to be set-theoretic; it had an intuitive content. The members of the set K were to be understood as so-called 'possible worlds;'[4] and a proposition (or sentence) was to be necessary iff true in all possible worlds.

---

[2] Presumably I meant the 1961 edition, current when I wrote this paper.
[3] He adds: "By implication, so much the worse for unquantified modal logic as well …" I myself assumed that modal logic is a proper subject only if quantified modal logic makes some sense. This isn't really the atmosphere today where much writing emphasizes only the propositional systems, which have application to computer science. See Goldblatt's history of modal logic (2006: 2), which explicitly says that he includes little about quantified modal logic.
[4] I have described the construction in Kripke (1959a) in the same way elsewhere, but strictly speaking, such a description is not correct. Possible worlds as such, do not appear as objects in the paper. But each possible world determines a model of a given formula, in which it is true or false. Many possible worlds may determine the same model of a given formula. The set K is the set of all models of a formula corresponding to some possible worlds. (And, of course, need not include all models, contrary to a misunderstanding I have seen.) A formula including modal operators is then defined as valid, if it is true in all elements of any non-empty set K.

Here the notions of necessity and possible world may be interpreted in several ways; perhaps the most common interpretation identified necessity with analyticity. (I will explain below how one might interpret 'possible world.') This interpretation of necessity as analyticity gives one an *applied* semantics for a system of modal logic; but this semantics is by no means the only applied semantics one might set up.

As Quine himself has pointed out (1976: 169), necessity might be construed more narrowly, as validity with respect to the logic of truth functions and quantifications and perhaps classes. Or it might be construed more liberally, as say some sort of physical necessity. But, if we are dealing with a single system of modal logic, all these alternative interpretations, giving different types of *applied* semantics, will nevertheless yield semantical notions having a common mathematical structure; and this mathematical structure – say the one set up in Kripke (1959a) – will be the *pure* semantics of the theory.

The distinction between *pure* and *applied* semantics can be made clear by analogy with the ordinary two-valued truth-functional propositional calculus. In *Principia*'s classic presentation, the calculus was developed syntactically, using a specified set of axioms and rules of inference. A semantical analysis of the system was discovered later by Post and Wittgenstein.[5] Formulae of sentential logic were to be assigned truth-values T or F according to certain rules. A formula was to be tautologous if it had the

(Actually, in the 1959a paper, I designate one element of K as G, the one corresponding to the real world. Then I define validity as truth in any such G. This doesn't matter for the purposes of the 1959a paper, since any G could be designated, but it is of some importance elsewhere.)

[5] *Principia Mathematica* already knew this semantics. See the interpretation of the "calculus" in terms of 0 and 1 in the first paragraph of section *4 (p. 120). They appear, however, not to regard this as the "intended" interpretation. (They presumably got this interpretation from the workers in Boolean algebra.)

The authors of *Principia Mathematica* already knew this semantics. See the interpretation of the "calculus" in terms of 0 and 1 in the first paragraph of section *4 (p. 120). Post and Wittgenstein appear, however, not to regard this as the "intended" interpretation. (Whitehead and Russell presumably got this interpretation from the workers in Boolean algebra.)

value T under all possible assignments; and the theorems of *Principia* demonstrably coincided with the tautologies. This was a *pure* semantics for set theory. T and F could be interpreted in whatever way we like; mathematically, we need consider only mappings of formulae into two distinct objects. Nor need we interpret the *Principia* '$\vee$' as disjunction or anything else; it need only obey an abstractly stated truth-table. This, then, is a *pure* semantics for two-valued logic; and it can be extended, as Gödel showed, to a pure semantics for extensional quantification theory.

But the pure semantics can be interpreted in several ways. One might interpret the sentential variables as ranging over English sentences, and T and F as ascribing truth and falsity, respectively, to the sentences. In this case, the truth-table for the connective '$\vee$' would require that it be interpreted as disjunction, the English inclusive truth-functional sense of 'or.' One might just as well take propositional variables as ranging over the same English sentences, but let T ascribe *falsehood* and F *truth*. The truth-table for '$\vee$' then would require that this connective be interpreted as *conjunction*, the English 'and.' *Principia* would be interpreted as asserting the self-contradictory character, not the truth, of its 'theorems.'

These two interpretations of *Principia*'s sentential calculus are radically different, yet they have a common mathematical structure revealed in the truth-table. The two different *applied* semantical readings correspond to a single *pure* semantics; and both the applied semantical systems are problematic in a way that the pure semantics is not. For as Quine's difficulties about indeterminacy show, the concept of truth is not clearly well-defined for arbitrary sentences of an arbitrary natural language; and it may not be a translation invariant. Again, when sentences become vague, the question of their truth becomes vague also; and in such cases as ethical sentences, we often don't know whether to speak of their truth-value.[6]

The answers to these questions about truth are by no means clear, but this does not mean that the pure semantics is unclear. Nor does it mean that there might not be interesting applied interpretations of *Principia*'s sentential logic, rigorously formulated, yet closely resembling in motivation the intuitive interpretation of T as applicable to true English sentences. Indeed, Tarski has given us a method, applicable to a large class of formal systems, of defining the concepts of truth and falsity; and the T and F may thus be interpreted as referring to truth and falsity, respectively, in a suitably formalized language. This applied semantics makes rigorous the notions of truth and falsity which were more vaguely used in interpreting English.

Quine's criticisms of modal logic have, as we have pointed out, taken several distinct forms. Some of Quine's published criticisms have seemed to indicate an expectation that any semantical treatment of modal logic, *pure* or *applied*, would fail, or else collapse modality. Such, for example, would seem to be the upshot of the discussion of modal logic in Quine (1960: 198). For although Quine asserts that his remarks are "predicated on the interpretation of modality as analyticity" (an *applied* semantics), the actual argument in that section, leading to a postulate that annihilates modal distinctions, would apply to any interpretation of necessity satisfying certain simple formal conditions. (Indeed the same applies to the so-called Morning-Star Paradox; it too can be stated formally on the basis of certain axioms governing descriptions.)

[6] I didn't really wish to commit myself here to Quine's views about the indeterminacy of translation. Also, as to vagueness, one possible view is that the truth predicate is vague as well. For ethical and other statements of value, the problem of whether they have truth-values is well known and I did not make any commitment about it.

Sometimes, on the other hand, Quine has maintained that quantified modal logic, though perhaps admissible, requires an intensional ontology which rules out individuals and classes; so Quine argues in the first edition of *LPV*.

Now, one of my primary purposes in Kripke (1959a) was implicitly to refute both of these claims: I presented a rigorous *pure* semantics for modality, with no special assumption on the character of the values of the variables. In the revised edition of *LPV*, Quine indeed explicitly acknowledges that limiting the values of one's variables to intensional entities is neither necessary nor sufficient for quantified modal logic. There is no explicit acknowledgement that a pure semantics for a quantified modal logic is indeed possible, but we shall interpret Quine's silence (in the revised *LPV*) on the point as a tacit acknowledgement of this fact also. [In order that the remaining issues be clearly defined, an explicit acknowledgment from Quine of this fact would be highly desirable.]

Although we interpret Quine as now at least tacitly acknowledging the possibility of a *pure* semantics for modal logic, we can suppose that he fears that the process of transforming it into an *applied* semantics may lead to special difficulties (at least in the intended interpretation), largely related to essentialism. The number nine, for example, is determinable by two conditions; it is the unique natural number x such that,

1. $\sqrt{x} + \sqrt{x} + \sqrt{x} = x \neq \sqrt{x}$

where only natural numbers are in the equation, as well as the unique natural number x such that,

2. There are exactly x planets.

Now (1), but not (2), must be taken as a necessary property of the number 9.[7] For if we wish to assert that,

3. $(\exists x)\ \mathbf{N}(8 < x < 10)$

where x ranges over natural numbers and **N** is necessity, it is clear that if 9 is that number it must be specified by a property like (1), not a property like (2). For it is by no means necessary that the number of planets is between 8 and 10. Further, it appears that any quantified modal logic must accept a distinction between the necessary and the contingent attributes of an object. Self-identity, for example, is surely a necessary property of every object: if we are permitted at all to apply necessity to an open sentence, surely $(x)\ \mathbf{N}(x = x)$ is true. But if p is any contingent statement, $x = x\ \&\ p$ will be true, but contingently true; and hence x has 'accidental,' as well as essential properties.[8] In "Three Grades," Quine expresses Aristotelian essentialism as saying "that you can have open sentences Fx and Gx such that,

4. $(\exists x)(\mathbf{N}Fx\ \&\ Gx\ \&\ {\sim}\mathbf{N}Gx)$." (1976: 174. Notation slightly revised.)

[7] As elsewhere, I have not changed Quine's example of 9 as the number of planets, though astronomers may reject it today. (It seems to me to be of clear philosophical significance that the astronomers could change their minds about the number of planets, but mathematicians cannot change their mind about what 9 is; though this has no significance for the present paper.)

[8] Actually, an interpretation of quantified modal logic is in fact possible that would restrict necessary properties to properties like self-identity that are obviously necessary independently of how an object is described, and are therefore unsusceptible to Quine's objections. This is not, of course, the interpretation I have come to prefer. See Fine (1978), section II.

So far, one cannot quarrel with Quine's assertions; but in one way, they are strangely formulated. Surely the objections to essentialism are not objections to the *truth* of a sentence of the form (4), but rather to the very *meaningfulness* of essential predication. Granted the *intelligibility* of this notion, it is not at all surprising that self-identity is an essential attribute of every object, or that $x = x \ \& \ p$ is contingent for contingent p. What is at issue is the question whether essential predication *makes sense*; i.e. whether,

5. $\mathbf{N}\mathrm{F}x$

ought to be well-formed as an open sentence. For (5) asserts that F is an essential attribute of the object x; and if this assertion is meaningful, it is by no means surprising that it is sometimes true. Thus I find it odd that Quine writes, in *LPV*, new edition, that Miss Barcan's system "hints" at essentialism through her theorem $(x, y)\ (x = y \supset \mathbf{N}(x = y))$[9] (*LPV*: 156). No theorems but simply the formation rules are needed to assure us of the trivial fact that all quantified modal logics are essentialist.

But this is a minor criticism. Let us return to the beginning. Quine, in both editions of *LPV*, differentiates between essentialism and analyticity thus:

[9] Actually, Barcan follows *Principia Mathematica* in defining identity as a second order notion. In *Principia* it is $(\mathrm{F})\ (\mathrm{F}x \supset \mathrm{F}y)$, and she also considers a variant notion with the material conditional replaced by a strict conditional. These notions are shown to be coextensive, or even strictly equivalent given S4. To get the necessity of identity out of this, one needs to use in addition what has been called the Barcan axiom.

I attribute the usual proof of the necessity of identity, taking identity as primitive, and which is much simpler, to Quine himself in Kripke (1959a).

> The Aristotelian notion of essence was the forerunner… of the modern notion of intension or meaning. For Aristotle it was essential in men to be rational, accidental to be two-legged. But there is an important difference between this attitude and the doctrine of meaning. From the latter point of view it may indeed be conceded (if only for the sake of argument) that rationality is involved in the meaning of the word 'man' while two-leggedness is not; but two-leggedness may at the same time be viewed as involved in the meaning of 'biped' while rationality is not. Thus from the point of view of the doctrine of meaning it makes no sense to say of the actual individual, who is at once a man and a biped, that his rationality is essential and his two-leggedness accidental or vice-versa. Things had essences, for Aristotle, but only linguistic forms have meanings. (*LPV*: 22)

As Quine points out in the revised *LPV*, the founders of modern modal logic (C. I. Lewis, Carnap) intended to base their system on a notion of analyticity, but instead they have based it on the much less clear notion of essentialism.[10] Quine argues that although in the doctrine of meaning it makes sense to say that *qua* the square of three nine is necessarily greater than seven, *qua* the number of planets its relation to 7 appears to be accidental. *Qua* mathematician one is necessarily rational but contingently two-legged; *qua* cyclist, the reverse. From the point of view of the doctrine of meaning, there is no distinction possible *qua* object, whether the man who is both cyclist and mathematician is essentially rational or not. Thus essentialism is baffling—more so than the modalities themselves. (Quine 1960: 109).

---

[10] Or, at least, they need to so if they intend modality to interact with quantification in the usual way. I am not certain why I included Carnap with C. I. Lewis. Lewis introduced modal logic (or really, originally, strict implication) into modern logic. I don't think he mentioned the combination with quantification. (Barcan, Carnap, and Church published the first quantified systems.)

The necessary distinctions are made out carefully and judiciously in Quine's "Three Grades of Modal Involvement." This paper, though written some years ago, comes remarkably close to Quine's present position; further, it is one of his most careful statements of the issue. He distinguishes three different degrees of acceptance of the notion of necessity. The first, or least degree of acceptance, is this: necessity is a semantical predicate applicable to names of statements. Thus we may say, abbreviating this predicate as 'Nec,'

6. Nec ('9 > 7')
7. Nec (Fermat's Last Theorem)

If necessity is analyticity, (6) says that '9 > 7' is analytic, while (7) says that Fermat's last theorem is analytic. But we might take necessity as a statement operator—the second degree of modal involvement. To do so we prefix necessity, now symbolized by N, to statements, not their names, obtaining, instead of (6) and (7), respectively:

8. **N**(9 >7)
9. **N**~(∃x, y, z, n) ($x > 0$ & $y > 0$ & $n > 2$ & $x^n + y^n = z^n$)

This permits, of course, iterated modal operators, since once we have attached a modal operator to a statement we can attach it again:

10. **NN**(9 > 7)
11. **N**~**N**~(9 > 7)

But still, necessity operators are inapplicable to open sentences.

The third grade, as we might guess, applies necessity to sentences, open and closed alike; thus it clearly involves one in essentialism. Now although in most publications Quine has interpreted necessity as analyticity, and predicated his remarks on this interpretation, it is clear that the problem of essentialism can be approached in a more general context, namely: what philosophical or logical differences are there among the various grades of modal involvement? What new assumptions are required for the transition from the first grade to the third?

Surprisingly enough, our answer will be: *none*. Anyone who accepts the first grade must accept the third; hence, in particular, anyone who accepts analyticity accepts essentialism also.[11] The latter cannot be more untenable than the former. And the founders of modal logic did not deceive themselves when they thought it could be based on analyticity alone. The core of this paper will be devoted to the detailed substantiation of this claim. The substantiation will not be 'philosophical,' but will be undertaken with mathematical rigor; so anyone who denies this denies a mathematical result. In particular, to set up a sharp opposition between the theory of essences and the theory of meaning is, in a sense, to deny a mathematical result. The claim is thus very strong; that my argument is as convincing as, say, the proof of the Cauchy integral theorem. However, after the argument has been presented, there will be some qualification, showing that there is some room for discussion left; and these remarks will be more nearly tentative.

[11] By "essentialism" I think I simply meant that modal operators can apply to open sentences, and that quantifying in to such modally operated open sentences is acceptable. No deeper philosophical doctrine was involved.

To proceed with the argument: Suppose we have a predicate 'is necessary,' applying to statements (of a language L already containing truth-functions) and hence attaching to names of statements, such that

12. Every necessary statement is true.

13. If p and $\ulcorner p \supset q \urcorner$ are necessary, so is q.

14. Every truth-functional tautology is necessary.

If analyticity is definable at all, it satisfies these conditions. But notice that a wide variety of semantical predicates satisfy them: logical truth, quantificational or set-theoretic validity, theoremhood in a particular formal system, necessity relative to certain premises, or even truth.

Now let us define a *maximal consistent set* of statements of L in the following manner: It is to be a set H of statements of L, such that,

15. For every closed sentence p of L, either p or $\ulcorner {\sim}p \urcorner$ is in H.

16. If $\ulcorner {\sim}p \urcorner$ is necessary, p is *not* in H.

17. If p and $\ulcorner p \supset q \urcorner$ are in H, so is q.

This notion, except for the reference to necessity, is a familiar device of modern extensional logic. Its modal connections have been insufficiently noticed; it corresponds to the modal notion of "possible world".[12] We now define an equivalence relation between sentences (closed) of L. Two statements, p and q, are *equivalent* if the

[12] However, Carnap's notion in *Meaning and Necessity* (see p. 9) of a "state-description," though not exactly the same as the one given here, is closely related. I should have noticed this fact.

biconditional ⌜$p \equiv q$⌝ is necessary. We say two such sentences determine the same *proposition*; we can then define a proposition as an equivalence class of statements. Propositions, thus defined, are not intensional enough to be objects of belief,[13] since analytically equivalent sentences determine the same proposition; but they are intensional enough for modal logic. Purists who object, reserving propositions as entities intensional enough to be believed, may replace my usage of 'proposition' by 'modal value of a proposition.' Now if p is a member of the set H, intuitively, we will say that p is *true* in the possible world H. A possible world, then, because of condition (15), must assign at least one truth-value to every statement; because of conditions (13), (16) and (17), it assigns at most one. A possible world is thus the totality of facts in that world. Thus equivalent statements, being true in precisely the same worlds, determine the same class K of worlds H; namely, p determines the set of worlds H such that p is true in H, and if ⌜$p \equiv q$⌝ is necessary, q is true in the same worlds as p. Conversely, if p and q express distinct propositions, then ⌜$p \equiv q$⌝ is not necessary, and a little mathematical detail, familiar to logicians and too long to be expounded for others, suffices to show that in this case there is a world H containing p but not q, or a world containing q but not p. So propositions determine in a one-one fashion classes of worlds; but it by no means follows that every class of worlds corresponds to a proposition, in the sense of proposition defined above—although, if it does correspond to a proposition, that proposition is uniquely determined. So we could have defined 'proposition' in a wider sense, thus: A proposition is a class of worlds; or, equivalently,

[13] In literature later than this paper, this assertion has not been universally accepted. There are those who may try to get out of the problem, or mitigate it. See, for example, D. K. Lewis (1979: section II, pp. 514-515 – and ignoring here his later modification for *de se* attitudes). Here even propositions in the broad sense of sets of worlds can be objects of arbitrary propositional attitudes, although Lewis acknowledges that this leads to problems.

it is a function assigning to each world a truth-value T or F. The propositions in the old sense will, in general, determine a subclass of the propositions in the new sense. In the new sense, whenever we are given a determination of truth-conditions, we have a proposition; or more exactly, when we are given a class K of worlds H, the class K can be interpreted as a proposition asserting that one of the worlds $H \in K$ is the real world. (N. B.: The set G of all the true statements satisfies the conditions (15)–(17); it can be called 'the real world.') Classes of worlds which correspond to no equivalence class of sentences might be said to be propositions "inexpressible" in the language L, but propositions nevertheless; for it is fully determined in what situations (worlds) they are true and in what situations they are false. In what follows, as long as possible, we will try to show how our concepts could be defined using *either* notion of proposition, the narrow one of an equivalence class of statements or the wider one of a set of worlds.

Now it is clear that truth-functional statement composition determines corresponding truth-functional operations on propositions as well. Let us use capital letters for propositions: P, Q, R, … Now if P and Q are any two propositions (in the wider sense), let them correspond to classes K_1 and K_2 of worlds; then $\ulcorner P \vee Q \urcorner$ is defined as the set-theoretic union $K_1 \cup K_2$; that is to say, $\ulcorner P \vee Q \urcorner$ is to be true in a world H if and only if either P or Q is true in H. If we look at a proposition in the narrower sense as an equivalence class of sentences, so that p is a sentence corresponding to the proposition P, and the same for q and Q, then we can stipulate that $\ulcorner P \vee Q \urcorner$ is the equivalence class determined by $\ulcorner p \vee q \urcorner$. We need only verify that if p is equivalent to p', and q to q', then $\ulcorner p \vee q \urcorner$ is equivalent to $\ulcorner p' \vee q' \urcorner$; and this verification is immediate. So in this manner all the truth-functions can be defined.

Now let P be any proposition in the wider sense. We define $\mathbf{N}(P)$ to be the class K of all worlds if P is the class of all worlds; otherwise $\mathbf{N}(P)$ is the empty set of worlds. This definition can be rephrased, more intuitively, thus: If P is true in all possible worlds, so is $\mathbf{N}(P)$; otherwise $\mathbf{N}(P)$ is false in all possible worlds. Alternatively, if we were using the narrow sense of proposition, so that P is the equivalence class determined by the sentence p, we could say: $\mathbf{N}(P)$ is the equivalence class of $\ulcorner p \vee {\sim}p \urcorner$ if p is necessary; otherwise $\mathbf{N}(P)$ is the equivalence class of $\ulcorner p \mathbin{\&} {\sim}p \urcorner$. It is easy to verify that the definition is independent of the choice of statement p. (Further, it should be verified that the two definitions, the one for the narrow sense and the one for the wider sense of proposition, coincide on the narrow sense; and this should also be verified for disjunction above.) In either definition of '**N**', **N** is a unary function mapping propositions into propositions; hence it can be iterated *ad libitum*, and combined with truth-functions at will.

Statements can be interpreted as denoting or meaning the propositions they determine; and thus the necessity operator can be attached to statements: If p determines the proposition P, we can use $\mathbf{N}({\sim}\mathbf{N}p)$ to denote the equivalence class $\mathbf{N}({\sim}\mathbf{N}P)$.

Thus the transition from the first to the second grade of modal involvement has been accomplished. In his paper on these grades, Quine remarks that "it is significant that in modal logic there has been some question as to just what might most suitably be postulated regarding such iteration" (169; note omitted). Here Quine understates his case; this question has long been one of the most perplexing and vaguest problems of modal logic. I have shown, in work announced in my abstract, "Semantical Analysis of Modal Logic" (1959b), that the step from the first to the second grade of modal involvement can be made in various ways, leading to alternative systems of iterated modalities. In fact, the major systems proposed in the literature can be obtained by some method at this stage.[14] For our present purposes it suffices to consider only one of these methods, the one used above; I mention that it leads to S5.

We have a modal propositional logic; how can we pass to a modal system with quantifiers? It is this step which Quine fears is fraught with danger and lurking metaphysical assumptions. Let us be very careful, then, that we do not introduce any new assumptions here; our assurance lies in the fact that, in this discussion we shall continue doing pure mathematics, and will not make philosophical remarks of any sort. The problem as stated above, is how to make sense of the locution,

5. **N**F*x*

Now, suppose we have a domain D of individuals, the values of the variables x of (5). Before we attempt to interpret (5), let us first attempt to interpret the simpler open sentence

18. F*x*.

'*x*' is supposed to be a free variable ranging over the domain D.

What sense, however, can we make of the letter 'F'? Classically, in the writings of Russell, 'F' was supposed to be a "propositional function", but this terminology has gone out of fashion. The term 'propositional function' was ambiguous; sometimes it

[14] I am not sure what I had in mind here. My modeling of the systems of modal propositional logic in terms of properties of a relation R between has now been published (Kripke 1963 and 1965), but I am not sure what I thought it had to do with the present construction.

meant 'attribute', sometimes 'open sentence' or 'statement matrix'.[15] But the uses of this notion in *Principia* make the proper interpretation clear; a propositional function must be interpreted as an attribute, not a statement matrix; and clearly attributes, whatever they may be, are within the province of modal logic. The 'F', then, in 'Fx', ascribes an *attribute* to the individual x. But what is an attribute?[16] The term "propositional function" gives us a clue long neglected in the literature. *An attribute is simply a mathematical function assigning a proposition to each element of the domain D.* Assume we have a particular propositional function F, and a particular individual x; then by definition F assigns a proposition P to x. If P is necessary (i.e., holds in all possible worlds), then we call F an "essential" attribute of x; if $\sim P$ is necessary, F is *excluded* by x; otherwise F is a *contingent* attribute of x. If Fx assigns P to x, $\mathbf{N}Fx$ assigns $\mathbf{N}$P; so, having made sense of (18), we have also made sense of (5).

One modification, to be sure, forces itself upon us when we introduce the quantifiers. We have defined Fx; how are we to define $(\exists x)$Fx? From the present point of view $(\exists x)$Fx can be defined as the (infinite) disjunction of all the propositions Fa, Fb, Fc, …, where a, b, etc. are the elements of D; i.e., if we consider all the propositions

[15] In reading Russell this way, I was following Quine. Perhaps some others have also read "propositional function" as meaning "open sentence." Today, I would not read Russell and *Principia* this way. A Russellian proposition was a structured object (see Kripke 2005: 237-8). A propositional function was a potential proposition with letters (blanks) to be filled by objects of the appropriate types. When filled in, it yielded a proposition. So it was a function of a special sort yielding propositions as values, but did not mean an arbitrary function from objects to propositions. It did not mean an "open sentence" (though I recall somewhere a loose passage in which it was called an "expression," where taken literally it was not a linguistic object at all). I do think that the concept of replacing objects in Russellian propositions was probably somewhat obscure, regarded non-linguistically.

(A complication is that later Russell came to repudiate propositions in this sense, and accepted only facts; but most of the interpretation of *Principia* goes in the old way.)

For the way I have been using the term "propositional function", see also "Quantified Modal Logic and Quine's Critique: Some Further Observations", fn 1.

[16] "Attribute" here is another piece of Quinean terminology. I don't know if it is the best way of describing 'propositional functions' (see the previous note and the discussion of what Quine's skepticism about 'quantifying in' might amount to if the theory is thought to be a theory of 'propositional functions'.

assigned to the elements of D by the propositional function F, $(\exists x)Fx$ is to be their disjunction. If the domain is finite, say a_1, …, a_n, the notion of the disjunction of all the propositions $Fa_1 \vee \ldots Fa_n$ is clear no matter which notion of proposition we adopt; for finite disjunctions have already been defined for either type of "proposition". If D is infinite, however, we need an infinite disjunction; and here is where the trouble begins. If we construe propositions in the wider sense, as classes of possible worlds, no difficulty ensues; for then the infinite disjunction of any set of propositions is simply the set-theoretic union of all the corresponding classes of worlds. The narrower version of proposition, however, according to which propositions are just equivalence classes of statements, yields no corresponding procedure without special assumptions on the nature of the language L. We might make these assumptions, but there is no need to do so; and we do not wish to complicate the present paper any further. So henceforth, by "proposition" we shall mean proposition in the wider sense; i.e., class of possible worlds.[17]

Existential quantification, then, comes to be explained as follows (if F is an attribute or propositional function, i.e., a function assigning a class of worlds to each member of D). Then, if D is the domain of this function, let X be its range; X is a set of classes of possible worlds. The union of all the classes in this set is what we mean by $(\exists x)Fx$. Finally, in terms of this, let us explain the notation $(\exists x)\mathbf{N}Fx$: If F assigns a proposition P to an object a, then **N**F assigns $\mathbf{N}P$ to the object a. Thus given an attribute F, we have defined an attribute $G = \mathbf{N}F$ expressing the necessity of the attribute F. To say $(\exists x)\mathbf{N}Fx$ is to say $(\exists x)Gx$; and the latter has been defined above.[18]

---

[17] This usage of 'proposition,' not just in the present construction, has become widespread.
[18] The definition can clearly extend to polyadic predicates.

So we have arrived at the third grade of modal involvement without recourse to any philosophical notion beyond that of a semantical predicate satisfying (12) - (14). Just as the step from the first to the second grade is by no means unique, so the step to the third grade is non-unique. But for the purposes of the present paper it suffices simply to show that there is at least *one* way of making this step.

Let us give some examples, to make the intuitive content of the preceding definitions clear: Consider the question how to make sense of the statement:

3. $(\exists x)\, \mathbf{N}(8 < x < 10)$

This statement asserts that the attribute of being between 8 and 10 is essential to some natural number. Now, how are we to define this attribute in terms of our formalism? Let D be the domain of non-negative natural numbers. To say $8 < x < 10$ is to say that $8 < x$ and $x < 10$. Let us analyze both of these. $8 < x$, or Fx, can be defined as the following proposition: The numbers 1, 2, …, 8 are to be mapped into the contradictory proposition (the empty set of worlds); the numbers 9, 10,… are to be mapped into the necessary proposition (the set of all possible worlds). Similarly we define $x < 10$ as necessary of 1,…, 9 and impossible of all other natural numbers. To adopt a different definition is simply to define a different attribute. Indeed, both of the attributes $8 < x$ and $x < 10$ are special cases of the binary attribute $x < y$, which is necessary of certain number pairs and impossible of others. When we assign the value 10 to y, we get the monadic attribute $x < 10$; and similarly for $8 < x$. If, in fact L already contains a notation for number theory, the attribute '<' can be defined in a more intuitive way: '<' is to assign to the pair of numbers m and n, the proposition determined by the sentence $0^{(m)} < 0^{(n)}$, where $0^{(m)}$ means 0 followed by m strokes. But it is noteworthy that the attribute '<' does not depend on the presence in L of any particular notation at all.

I do not mean to claim that the method I have used to introduce quantified modality is the only conceivable method of doing so, or the best, or that other methods may not lead to peculiar philosophical problems. I only claim that I have given *one* method of setting up quantified modal logic; and that this method requires only the existence of a semantical predicate satisfying (12) - (13). Does this show that essentialism is after all *not* involved in quantified modal logic, that essentialism can be defined in terms of analyticity, or that analyticity involves as *bad* a metaphysics as essentialism? The option is unreal.[19] In this paper, we have shown the following simple fact: If analyticity is well defined, so is quantified modal logic. And even if analyticity is *not* well defined, there are other necessity predicates, such as theoremhood in a specified logical system. The modal systems built on these predicates are interesting even if the classical applied semantics, in terms of analyticity, is not.

Having made these remarks, I qualify them as follows: The procedure adopted, though satisfactory for all purposes of quantified modal logic, is unsatisfactory if one wishes to make a further demand. The language L may itself contain quantifiers, ranging over a universe U; what is the relation between these quantifiers and the quantifiers introduced above, over a domain D? The answer: there need be none. It may even be necessary in L that U contains at most three elements; this by no means prevents us, in our previous construction, from using a domain D with four or more

---

[19] Here I was no doubt imitating *Word and Object*, p. 265 (see the remarks on physicalism and solidity). Actually, in the technical sense that I was using "essentialism", of course no special metaphysics is involved.

elements. The question then, arises: Can one introduce quantified modality in such a way that the quantifiers introduced have a "natural" connection with the quantifiers in L? A discussion of this (vaguely formulated) problem would prolong this paper to an inordinate extent. Let me simply assert that in the analysis of this problem, difficulties of "essentialism" very similar to those suggested by Quine do indeed arise. But these difficulties are at most objections to one way of introducing quantified modalities; they do not apply to the method sketched above.

Let us conclude with certain remarks on the intuitive problem of essentialism. Basically the problem is as follows: What attributes must an individual, say Jones, have in order to qualify as Jones? Which of his attributes can vary from possible world to possible world, and which must remain? The former are contingent, the latter necessary. Presumably, for example, though Jones is a mathematician, he would still be Jones; he would be Jones had he been born a day earlier; so both his mathematical profession and his birthday are contingent attributes. But could we imagine of Jones that he was not a human being but a planet? If not, being a non-planet is essential to Jones, and we would count this attribute as essential. The intuitive question of essentialism boils down to this: How does an individual preserve his identity from one possible world to another? In the case of human beings, we have only vague criteria of identity and identification, so the question is vague (though not, I think, intuitively *meaningless*). In the case of natural numbers, on the other hand, we use precise criteria of identity: the identity of a natural number depends on its position in the natural number series, and nothing else. Hence (1), but not (2), is intuitively an essential attribute of the number 9; for it is (1), but not (2) that follows from 9's position in the series. That this account of essentialism agrees with our intuitive notions is readily confirmed by the fact that no one, intuitively, could ever take the reverse line of declaring (2), and not (1), an essential attribute. If the world changed, the number of planets would be different; but we would say, not that a particular natural number suddenly changed its position in the series, but that there were more or fewer planets.[20]

In view of these facts, it is somewhat odd that Quine counts the distinction between a fleeting and an enduring attribute as clearer than that of a necessary or contingent one (Quine 1960: 199). For, as Prior has pointed out (and made this point the subject of an entire book in Prior 1957), if we try to make a tense logic for temporal (not eternal) sentences, the instants of time play a role formally quite analogous to the possible worlds. A sentence which is always true counts as necessary. Again a problem of essentialism arises: how does an object preserve its identity from one instant to another? Imagine, e.g., a building built with 10,000 bricks, and then torn down brick by brick. When one brick is torn down, the building remains; when 8000 are lost, it does not. For what *n* does possession of at least *n* bricks become an enduring attribute of the building, one which remains a property of the building as long as the building itself endures? Such questions can be set aside by dealing only with objects defined independently of time, e.g., the four-dimensional "process-things," abstract entities, and the like; and similarly, one can set aside questions of essentialism by introducing the domain of objects D independently of the possible

[20] In this paragraph we can see how my views have changed in *Naming and Necessity*. In 1961/2, when this paper was written, I had at least some of the intuitions about essential and accidental properties of objects that I came to have later, though not all. However, I thought that what properties were essential depended on "criteria of identity [of individuals] across possible worlds". The idea that such criteria are needed is of course, emphatically repudiated in the later work. I also, as the present paper reads, did not already have in mind any fundamental distinction between *a priori* and necessary truths, whereas this distinction is basic in *Naming and Necessity*. Nor did I distinguish these notions from analyticity.

worlds. Essentially it was this expedient which was followed above, when possible worlds were certain classes of statements of a language L. But the doctrine that an essentialism of eternal attributes is intrinsically clearer than the essentialism of modality, is one for which I can find little justification.[21]

I am well aware of the expository inadequacy of these final remarks, and of their incompleteness, forced on us for reasons of space. Let us then recall the central thesis of this paper: that quantified modal logic can be developed on the slender basis of a single semantical necessity predicate. The development of this thesis, prompted by Quine's criticisms, may (so it is hoped), clarify the foundations of modal logic. Thus in the same spirit that one might say that Berkeley, who by his searching criticism of the calculus prompted men more sympathetic to it than he to clarify its foundations, made a substantial contribution to the calculus, so I think it may be said that Quine is one of the principal contributors to modern modal logic.[22]

Saul A. Kripke

The Saul Kripke Center and the Graduate Center of the City University of New York

[21] I would not view these matters the same way today either, though maybe the example of the building may remain.

[22] I would like to thank Gary Ostertag, Rosemary Twomey, and especially Romina Padró for their help in producing the present version. This paper has been completed with support from the Saul A. Kripke Center at The City University of New York, Graduate Center.

References

Carnap, R. (1947). *Meaning and Necessity*. University of Chicago Press.

Fine, K. (1978). "Model theory for modal logic – Part II. The elimination of *de re* modality." *Journal of philosophical logic*, vol. 7: 125-156.

Goldblatt, R. (2006). Mathematical Modal Logic: A View of Its Evolution. In *Handbook of the History of Logic Volume 7*, Gabbey, D. and Woods, J (eds.), North Holland Publishing Co.: 1-98. (Check)

Kripke, S. (1959a). A Completeness Theorem in Modal Logic. *Journal of Symbolic Logic* 24: 1-14.

Kripke, S. (1959b). Semantical Analysis of Modal Logic. (Abstract) *Journal of Symbolic Logic* 24: 323–324.

Kripke, S. (1963). Semantical Analysis of Modal Logic I. Normal Propositional Calculi. *Zeitschrift fur mathematische Logik und Grundlagen der Mathematik* 9: 67-96.

Kripke, S. (1965). Semantical Analysis of Modal Logic II. Non-Normal Modal Propositional Calculi. In *The Theory of Models* (Proceedings of the 1963 International Symposium at Berkeley), Addison, J. W., Henkin, L., and Tarski, A. (eds.), North Holland Publishing Co.: 206-220.

Kripke, S. (2005). Russell's Notion of Scope. *Mind* 114: 1005-37. Reprinted in Kripke (2011): 225-253.

Kripke, S. (2011). *Philosophical Troubles. Collected Papers Volume I.* Oxford University Press.

Lewis, D. (1979). Attitudes *De Dicto* and *De Se*. *Philosophical Review* 88: 513-43. Reprinted in Lewis (1983).

Lewis, D. (1983). *Philosophical Papers, Volume I*. Oxford University Press.

Prior, A. (1957). *Time and Modality*. Oxford University Press.

Quine, W.V.O. (1953). *From a Logical Point of View*. Harvard University Press. Second Edition, 1961.

Quine, W.V.O. (1960). *Word and Object*. MIT Press.

Quine, W. V. O. (1976). "Three Grades of Modal Involvement." *The Ways of Paradox*, Second Edition. Cambridge: Harvard University Press, pp. 158-76. Originally published in *Proceedings of the XIth International Congress of Philosophy*, vol. 1 (Brussels: North Holland), 1953.

Whitehead, A. N., and B. Russell. (1910, 1912, 1913). *Principia Mathematica, 3 Vols*, Cambridge, Cambridge University Press. Second edition, 1925 (Vol. 1), 1927 (Vols. 2, 3).