MODEL THEORY AND MODAL LOGIC

John L. POLLOCK

The interpretation of first-order logic in terms of models is well known. Recently attempts have been made to extend the model-theoretic results obtained for first-order logic to modal logic (1). The best known of such attempts are those of Kripke [1], Hintikka [3] and [4], Carnap [5], and Montague [6]. These involve modifications of the concept of a model. Although the modifications lead to interpretations which appear superficially very different, they can be reformulated in a way which shows that they are really very similar to one another. It is the contention of this paper that these attempts to formulate the semantics of modal logic all suffer from an apparently irreparable difficulty.

Let us suppose that our object language L is an ordinary first-order language, without predicate constants or individual constants, and augmented with the new logical constant $\square$ (2). Let us define a *semi-model* on D to be an ordered pair $\langle D,G \rangle$ where D is a non-empty set, and G a function mapping individual variables *onto* D, and for each $n > 0$, mapping R_n, the set of n-ary relation symbols, onto $S(D^n)$, the set of all sets of n-tuples of elements of D. If M is the semi-model $\langle D,G \rangle$, let $M^*=G$, and $M_*=D$. Now we can define a *model* to be an ordered pair $\langle G,K \rangle$ where K is a set of semi-models, $G \varepsilon K$, and for every $H \varepsilon K$, $H_* = G_*$. Let us say that $\langle G,K \rangle$ is a *prime* model if and only if K is the set of *all* semi-models H such that $H_* = G_*$.

Now we can define truth in a model as follows, where *val* is a function mapping the set of sentences onto $\{T,F\}$:

i) If $f \varepsilon R^n$ and $x_1,...,x_n$ are individual variables, then $val_{(G,K)} f x_1...x_n = T$ if and only if $\langle G^*(x_1),...,G^*(x_n) \rangle \ \varepsilon \ G^*(f)$;

(1) By 'modal logic' I mean the logic of *logical* necessity, as opposed to physical necessity, epistemic necessity, deontic necessity, etc.

(2) The symbols we write will be the names of symbols of the object language rather than the symbols themselves, so that, e.g., 'pvq' is not a sentence, but the name of a sentence.

ii) $val_{(G,K)} \sim p = T$ if and only if $val_{(G,K)} p = F$;
iii) $val_{(G,K)} p \& q = T$ if and only if $val_{(G,K)} p = T$ and $val_{(G,K)} q = T$;
iv) $val_{(G,K)} (x)fx = T$ if and only if for every individual variable y, $val_{(G,K)} fy = T$;
v) $val_{(G,K)} \Box p = T$ if and only if for every $H \varepsilon K$, $val_{(H,K)} p = T$.

Now let us say that a sentence is KH-valid if and only if it is true in every model, and CM-valid if and only if it is true in every prime model. It is not difficult to see that the definitions of validity given by Kripke and Hintikka (3) correspond to KH-validity, and the definitions given by Carnap and Montague correspond to CM-validity (4). These authors differ in their treatment of individual constants and identity, but that need not concern us here with our somewhat impoverished language L. These authors all give more or less the same intuitive explanation of their concepts of validity. The idea is that a model $\langle G,K \rangle$ can be interpreted as follows. K is the set of "possible worlds", and G is the actual world. Truth in a possible world $H \varepsilon K$ for a non-modal sentence is defined just as in ordinary first-order logic. Then to say that $\Box p$ is true in G is just to say that p is true in all possible worlds, i.e., true in every $H \varepsilon K$.

The difference between KH-validity and CM-validity is that in the former case, any set of semi-models on G^* is said to generate a counterpart of the set of all possible worlds, whereas in the latter case, only the set of *all* semi-models on G^* is taken as generating a counterpart of the set of all possible worlds. An immediate intuitive difficulty arises in connection with both of these methods of interpretation. Intuitively, there is only *one* set of all possible worlds. Why then are we allowed to assume that all of these (prime) models mirror the notion of the set of all possible worlds? A sort of an answer can be given by saying that in logic we are only concerned with the formal structure of the set of all possible worlds, and that this formal structure is obtained by requiring truth in all (prime)

(3) Here I am only talking about that system of Hintikka's which is intended to be an interpretation of *logical* necessity (in which alternativeness is an equivalence relation), and which does not employ his restrictions for non-referring individual constants. Those latter restrictions are irrelevant for our present purposes.

(4) See Montague's reduction of Carnap's formulation to this form, in [6].

models. I find that answer somewhat obscure, and I think that that obscurity is the source of the purely formal difficulties that I will point out shortly. But let us suppose for the moment that this answer is unobjectionable. This answer does not clearly determine whether we should consider all models, or just all prime models. I think, however, that it does make any intermediate alternative, which would consider more than just the prime models, but not all models, highly implausible. If this type of approach is to work at all, either KH-validity or CM-validity must succeed in formalizing our intuitive concept of a truth of the logic of *L*. It will now be shown that that is not the case for either KH-validity or CM-validity.

The defects in KH-validity and CM-validity are illustrated by noting that $\sim \Box[(\exists x)fx \;\&\; (\exists x)\sim fx]$ *is not* KH-valid (although it is CM-valid), and that $\sim \Box(x)fx$ is CM-valid (although it is not KH-valid). To see that these are in fact defects, let us reflect on what it means to say that a sentence is a truth of the logic of *L*. I would propose the following: A sentence *p* is a truth of the logic of *L* if and only if every interpretation of *p* is necessarily true ([5]). This would seem to be just what we mean by a truth of logic. Given this definition, it is clear that $\sim \Box(x)fx$ is not a truth of logic. There will be interpretations of the predicate letter *f* which will make $(x)fx$ necessarily true, and so $\sim \Box(x)fx$ will be false under some interpretations of *f*. Thus any system of interpretation which would make $\sim \Box(x)fx$ valid must fail to formalize our intuitive concept of a truth of logic.

Given our definition of a truth of logic, it also seems that the sentence $\sim \Box[(\exists x)fx \;\&\; (\exists x)\sim fx]$ is a truth of logic, and so should be valid. This is because, given any sentence *p*, if in order for *p* to be true there must be at least *n* things in the universe ($n > 1$), then *p* cannot be necessarily true. This is presupposed by our ordinary interpretation of first-order logic. But now, for any interpretation of *f*, the sentence $(\exists x)fx \;\&\; (\exists x)\sim fx$ requires for its truth that there be at least two things in the universe. Thus no interpretation of it is necessarily true. That is, every interpretation of the sentence $\sim \Box[(\exists x)fx \;\&\; (\exists x)\sim fx]$ is true. But assuming that

([5]) If this is not obvious, I would like to refer the reader to [7].

$$(\sim \Box p \supset \Box \sim \Box p)$$

is a truth of logic (and this is accepted by all of the above authors), it follows that every interpretation of

$$\sim \Box [(\exists x)fx \,\&\, (\exists x) \sim fx]$$

is necessarily true, and hence that

$$\sim \Box [(\exists x)fx \,\&\, (\exists x) \sim fx]$$

is a truth of logic.

Now we must prove the following:

(I) $\sim \Box[(\exists x)fx \,\&\, (\exists x) \sim fx]$ is not KH-valid.

(II) $\sim \Box(x)fx$ is CM-valid.

First let us prove (I). Choose any semi-model G such that $G^* = \geq$ 2. For some $f\, \varepsilon R_1$, G^*, and $G^*(f) \neq 0$.

Then $\langle G,\{G\}\rangle$ is a model, and $val_{(G,\{G\})}[(\exists x)fx \,\&\, (\exists x) \sim fx] = T$.

Then for every H $\varepsilon\{G\}$, $val_{(H,\{G\})}[(\exists x)fx \,\&\, (\exists x) \sim fx]$ T.

But then $val_{(G,\{G\})} \Box[(\exists x)fx \,\&\, (\exists x) \sim fx] = T$.

Therefore, $\sim \Box[(\exists x)fx \,\&\, (\exists x) \sim fx]$ is not KH-valid.

Now let us prove (II). Consider an arbitrary prime model $\langle G,K \rangle$. For every $f\, \varepsilon R_I$, there is an $H\, \varepsilon K$ such that $H^*(f) \neq G^*$ (as $G^* \neq 0$).

Then $val_{(H,K)}\, (x)fx = F$.

But then $val_{(G,K)} \Box(x)fx = F$, and $val_{(G,K)} \sim \Box(x)fx = T$.

As this is true for arbitrary $\langle G,K\rangle$, it follows that $\sim \Box(x)fx$ is CM-vaiid.

Thus neither KH-validity nor CM-validity succeeds in formalizing our intuitive concept of a truth of the logic of L (6). I am inclined to ascribe the defects of these interpretations to the fact that they relativize the concept of a possible world to a certain set K, and then vary K for their definition of validity. There can really only be *one* set of all possible worlds, and we must find a semantical interpretation which conforms to this fact.

State University of New York at Buffalo John L. POLLOCK.

(6) In [2], Kripke presents a definition of validity which is similar to but not equivalent to KH- validity. However, anything valid in that sense is also KH-valid, and so the above difficulty with KH-validity is also a difficulty with that concept of validity.

REFERENCES

[1] Saul A.KRIPKE, „A Completeness Theorem in Modal Logic", *The Journal of Symbolic Logic*, vol. 24 (1959), pp. 1-15.

[2] Saul A. KRIPKE, „Semantical Considerations on Modal Logic", *Acta Philosophica Fennica*, Fasc. XVI (1963), pp. 83-94.

[3] Jaakko HINTIKKA, „Modality and Quantification", *Theoria*, (Lund), vol.27 (1961), pp. 119-128.

[4] Jaakko HINTIKKA, "The Modes of Modality" *Acta Philosophica Fennica*, Fasc. XVI (1963), pp. 65-81.

[5] Rudolph CARNAP, "Quantification and Modalities", *The Journal of Symbolic Logic*, vol. 11 (1946), pp. 33-64.

[6] Richard MONTAGUE, "Logical Necessity, Physical Necessity, Ethics, and Quantifiers", *Inquiry*, vol. 3 (1960), pp. 259-269.

[7] John L. POLLOCK, "Logical Validity in Modal Logic", *The Monist*, vol. 51 No. 1 (*The Present Situation in Philosophic Logic*).