A Philosophical Conception of Propositional Modal Logic*

Edward N. Zalta
Center for the Study of Language and Information
Stanford University

The definitions of propositional modal logic are traditionally formulated in the following way.[1] First, a formal language is defined, usually with atomic formulas $p, q, \ldots$ and complex formulas involving the connectives $\neg$, $\rightarrow$ and the $\Box$ operator. Models $\mathbf{M}$ for this language are then defined as triples $\langle \mathbf{W}, \mathbf{R}, \mathbf{V} \rangle$ in which $\mathbf{W}$ is a nonempty set of worlds, $\mathbf{R}$ is an accessibility relation, and $\mathbf{V}$ is a valuation function that maps each atomic sentence of the formal language to a set of worlds. Truth (at a world, in a model), validity, and logical consequence are then defined as semantic properties of, or relations among, sentences of the language. Finally, a proof theory is developed so that the consequences of (sets of) sentences may be derived. In recent developments of this proof theory, rules of inference are conceived as relations between sentences, and a *logic* Σ is defined to be a set of sentences closed under certain rules. With such a framework of definitions, modal logicians then investigate metatheoretic questions, such as whether sentences valid in certain models are theorems (i.e., elements) of certain logics, etc.

From a philosophical point of view, however, this traditional conception of propositional modal logic has some unsatisfactory features, and in this paper, we shall argue for a different conception. Under the conception for which we argue, propositional modal logic ('ML') is concerned with propositions rather than sentences. A domain $\mathbf{P}$ of structured propositions (i.e., structured meanings) is distinguished, and models $\mathbf{M}$ represent the change in truth value of these propositions at various possible worlds. The notions of truth and validity are construed as properties of propositions. The notions of a rule of inference and a logic are construed, respectively, as a relation among propositions and a set of propositions closed under certain rules. Of course, in addition to this, a language may be defined. The interpretation of the language consists of a function that maps the sentences of the language to propositions. Given such an interpretation, one may define secondary notions of truth, validity, rule of inference, and logic, all of which apply to sentences.

In what follows, we precisely define this philosophical conception of ML and describe its advantages. In Section 1, however, we begin by describing the unsatisfactory features of the traditional conception to which we alluded above. This serves to motivate Section 2, in which we frame the precise definitions that embody the philosophical conception. In Section 3, we compare and contrast these definitions with the traditional ones. Finally, in Section 4, we discuss one of the more interesting consequences of the philosophical conception, namely, that modal contexts are not, strictly speaking, intensional contexts. The paper concludes with some final observations.

The philosophical conception of ML is grounded by a precise theory of propositions, and in particular, by the theory of propositions derivable from the theory of relations developed in Zalta [1983], [1988a], and [1993].[2] On this theory, propositions are defined as 0-place relations that are governed by: (a) a derived abstraction schema that comprehends the domain of propositions, and (b) theoretically-defined identity conditions which permit necessarily equivalent propositions (i.e., propositions p and q such that $\Box(p \leftrightarrow q)$) to be distinct. Since this theory has been published elsewhere, we shall not go through the details here. Readers unfamiliar with this background theory of propostions should not be at a loss, how-

*Published in *Philosophical Topics*, 21/2 (1993): 263-281. (This was a volume on Philosophy of Logic, C. Hill (ed.), Fayettesville: University of Arkansas Press.) Some of the ideas in this paper were conceived while writing the first chapter of an introductory text in modal logic (in progress). I would like to thank Bernie Linsky, Greg O'Hair, John Etchemendy, Chris Menzel, Kees van Deemter, Johan van Benthem, Uwe Mönnich, and Tony Anderson for their insightful contributions to discussions of the topics contained herein. I am also indebted to the Center for the Study of Language and Information for providing me with office space for the 1992-93 academic year.

[1]See, for example, Kripke [1963], Hughes and Cresswell [1968], Segerberg [1971], Montague [1974], Lemmon and Scott [1977], Chellas [1980], Dowty, Wall, and Peters [1981], Bull and Segerberg [1984], Goldblatt [1987], and van Benthem [1988]. The following description ignores trivial and inessential differences among these various formulations.

[2]The theories of propositions developed by Bealer [1982] and Menzel [1986] would serve equally well to ground our conception.

ever, since those details of the theory that are required for developing the philosophical conception of ML will be introduced at the appropriate time. For the rest, the mere fact that the background theory of propositions exists and precisely grounds the conception should suffice.

§1: Motivation for the Philosophical Conception

One, but by no means the only, reason for developing a new conception of ML is to resolve an ambiguity about what traditionally conceived models represent. To help us understand what the ambiguity is, recall that the truth of a sentence depends both on what it means and the way the world is. The truth of the sentence 'Snow is white', for example, depends both on the fact that it means that snow is white and on the fact that the substance snow has the property of being white. A natural way to distinguish these two ingredients for truth is to point out, counterfactually, that 'Snow is white' would have turned out false: (a) if it had meant that grass is purple, or (b) if the substance snow had had the property of being green.[3] In the first alternative, the sentence 'Snow is white' would have been false *if our language had been different*, that is, if the words 'snow' and 'white' had acquired different meanings. In the second alternative, the sentence would have been false *if our world had been different*, that is, if snow had been some other color.

Ordinarily when we express ourselves using modal language, we are concerned with the second rather than the first alternative. When we ask, might snow have failed to be white, we are not asking about whether the sentence 'Snow is white' could have expressed a falsehood, but rather whether the substance snow might have been some other color. When we assert that snow might not have been white, we are not asserting, even in part, that the sentence 'snow is white' might have meant something else. Instead, we are asserting something about the modal nature of (things in) the world.

The traditional models of the modal language of ML, however, do not distinguish these two alternatives. The $\mathbf{V}$ function of models maps the atomic sentences of the language to sets of worlds. It essentially evaluates

[3] Much of what I have to say in the next few paragraphs has been influenced by the ideas in Etchemendy [1990]. I've also benefited from reading Menzel [1990]. While reflecting on the ideas in these works, it occurred to me that there was a natural way to modify the semantic development in Zalta [1983] and [1988a] to make it clear that ordinary sentences do not change meanings when evaluated at other worlds.

them as true or false in counterfactual circumstances. An atomic sentence 'p' is true in all those worlds $\mathbf{w}$ such that $\mathbf{w} \in \mathbf{V}(p)$, and false in the others. But there are two ways to interpret the fact that $\mathbf{w}_1 \in \mathbf{V}(p)$, say, but $\mathbf{w}_2 \notin \mathbf{V}(p)$; in other words, the models do not distinguish between the following two interpretations for the fact that 'p' is true in $\mathbf{w}_1$ but not in $\mathbf{w}_2$:

(a) 'p' means one thing in $\mathbf{w}_1$ and means something else in $\mathbf{w}_2$, and moreover, what it means in $\mathbf{w}_1$ is true in $\mathbf{w}_1$ and what it means in $\mathbf{w}_2$ is false in $\mathbf{w}_2$

(b) 'p' has the same meaning in both $\mathbf{w}_1$ and $\mathbf{w}_2$, but what it means is true in $\mathbf{w}_1$ and false in $\mathbf{w}_2$

For all that the models tell us, it might well be that sentences such as '$\Diamond\neg p$' are true in virtue of circumstances such as (a). Yet this is something that ought to be ruled out.

It is no good to argue that formal sentences such as 'p' are presupposed to have a meaning (just as 'Snow is white' already has a meaning), that our models are therefore defined for an *interpreted* language, and that they represent the truth conditions of already meaningful sentences. For there is nothing in the traditional conception that tells us wherein the meaningfulness of the formal sentences lies. Of course, we might take the Montagovian intension of 'p' (i.e., the function that maps a world $\mathbf{w}$ to the truth value $\mathbf{T}$ if $\mathbf{w} \in \mathbf{V}(\mathbf{p})$, and maps $\mathbf{w}$ to the truth value $\mathbf{F}$ otherwise) as the meaning of 'p', and introduce some changes into the basic definition of a model, to reflect that we are trying to model the second of the two alternatives described above. But that explicitly identifies the meaning of all necessarily equivalent propositions, contrary to intuition. And besides, there is a better way.

To resolve this ambiguity, we shall need to distinguish sentences from the propositions they denote (or express, or signify). To develop this distinction, let us review some of the details of the theory of propositions that grounds our conception. The theory of propositions in Zalta [1983], [1988a], and [1993] is couched in a formal language with a formal semantics. This semantics gives us a precise way to talk unambiguously about the proposition that a sentence denotes. Consider a simple English sentence such as 'John is happy', which is represented in the object language by the formal sentence 'Hj'. The formal sentence is not only an atomic formula, but also a 0-place (complex) term. In virtue of being a term, 'Hj' receives a denotation in addition to the truth conditions it receives in virtue of being a formula. A recursively defined denotation function identifies the denotation of the complex term 'Hj' on the basis of the denotations of the component terms. Now let **H** be the property denoted by the predicate 'H', and let **j** be the object denoted by the name 'j'. Then, the denotation function identifies the denotation of 'Hj' as a structured proposition (i.e., 0-place relation) which can be formally described in terms of an algebraic logical operation $\mathbf{PLUG}_i$ which, among other things, harnesses a property (i.e., 1-place relation) and an object into a proposition.[4] Specifically, the denotation function has the following consequence:

'Hj' denotes $\mathbf{PLUG}_1(\mathbf{H},\mathbf{j})$.

So the denotation of the formal sentence 'Hj' is the structured proposition having the denotation of the object term 'j' *plugged* into the first (and only) place of the property denoted by the predicate 'H'.

There are three important features of this analysis worth noting. First, the property **H** denoted by the predicate 'H' is *not* to be conceived as a set-theoretic entity of any kind, but is rather to be construed as a primitive entity. For those who want to know more about such entities, my previously cited works also contain a general theory of properties, including both comprehension and identity principles for that domain.[5] Second, the denotation function is not relativized to a world. In particular, the denotation of the predicate 'H' is simply the property **H**, and this does not change from world to world. Of course, the *extension* of the property **H** may vary from world to world, but the meaning of the predicate (i.e., the property it denotes) is fixed. Similarly, the denotation of the whole sentence 'Hj' is independent of the worlds. However, the proposition it denotes, $\mathbf{PLUG}_1(\mathbf{H}, \mathbf{j})$, has an extension (a truth value) that varies from world to world. Third, constraints on the extension function for properties and propositions require that the extension of $\mathbf{PLUG}_1(\mathbf{H}, \mathbf{j})$ at a world **w** be the truth value **T** iff **j** is an element of the extension of the property **H** at **w**.[6]

Now, with this formal analysis of atomic propositions in mind, let us return to an example of a simple English sentence. Let '**being happy**' be a semantic name of the property denoted by the predicate 'is happy' and let '**John**' be the semantic name of the object denoted by the name 'John'. Then, by analogy with our formal language and formal semantics, we may say with some precision:

'John is happy' denotes $\mathbf{PLUG}_1(\mathbf{being\ happy}, \mathbf{John})$

Such claims, in which a sentence is distinguished from the proposition it signifies, should now be clear enough for us to proceed with our discussion of the ambiguity inherent in the traditional conception of models. For historical reasons, however, let us revert to the example 'Snow is white', despite the problems inherent in the analysis of mass terms such as 'snow'. Let us suppose that claims such as the following are clear enough:

'Snow is white' denotes **Snow is white**.

'Grass is purple' denotes **Grass is purple**.

If greater clarity is needed, the reader may now substitute examples such as 'John is happy'.

[4] **PLUG** was developed independently in Bealer [1979] (who calls it 'pred') and in Parsons [1980], and used in McMichael and Zalta [1980], Bealer [1982], Zalta [1983] and Menzel [1986]. The subscript 'i' in '$\mathbf{PLUG}_i$' signifies the place in the relation into which the object is plugged. So $\mathbf{PLUG}_2$, for example, could be used to plug an object into the second place of a 2-place relation, or into the second place of a 3-place relation, etc.

[5] In the cited works, the reader will find motivation for and explanation of the following comprehension and identity principles governing properties:

$\exists F\Box\forall x(Fx \leftrightarrow \varphi)$, where φ has no free Fs, no encoding formulas, and no quantifiers binding relation variables.

$F\!=\!G =_{df} \Box\forall x(xF \leftrightarrow xG)$

These principles are developed in terms of the notion of an object x *encoding* a property F (symbolically: xF), a notion which is systematized in the cited works. The theory of propositions can be stated in terms of the property theory:

$\exists p\Box(p \leftrightarrow \varphi)$, where φ has no free ps, no encoding formulas, and no quantifiers binding relation variables. (footnote continues on next page)

$p\!=\!q =_{df} [\lambda y\, p] = [\lambda y\, q]$

In these principles, 'p' and 'q' range over propositions (i.e., 0-place relations). In the identity principle, $[\lambda y\, p]$ and $[\lambda y\, q]$ denote 'propositional properties' (properties that things exemplify in virtue of propositions being true or false). So the identity principle says that propositions p and q are identical iff the property *being such that* p is identical to the property of *being such that* q. So the identity of propositions is defined in terms of the identity of properties. Compare Myhill [1963].

[6] The domains containing **j** and **H** are fixed and independent of the worlds, and so these entities 'exist' at all other worlds. This, of course, is not to say that the object **j** has a location in spacetime at all worlds.

So given the theoretical distinction between a sentence and the proposition it expresses, there is a natural way to describe in theoretical terms the two kinds of counterfactual circumstances in which 'Snow is white' would have turned out false. 'Snow is white' would have been false: (a) if 'Snow is white' had denoted a false proposition, say **Grass is purple**, instead of denoting **Snow is white**, or (b) if the proposition 'Snow is white' in fact denotes, namely, **Snow is white**, had been false. In circumstances such as (a), 'Snow is white' would have had a different meaning (the language would have been different). We don't want the models of modal logic to leave open the question of whether circumstances such as (a) make the sentence 'Snow might not have been white' true. We can rule out this question by making sure that the proposition that a sentence denotes is independent of the worlds in a model, so that models cannot be interpreted as assigning some other proposition to a sentence at other possible worlds. In circumstances such as (b), the world would have been different, since if the proposition **Snow is white** had been false, the substance snow would not be in the extension of the property of being white. We want our models of modal logic to reflect that circumstances such as (b) make the sentence 'Snow might not have been white' true. We can do this by defining models so that the **V** function (or something equivalent to it) assigns truth values to *propositions* at each possible world, rather than to sentences. Under such a conception, models need not say which proposition a sentence denotes (i.e., they need not tell us about the meaning or interpretation of the language); rather, they simply describe the nature of other possible worlds by recording what truth values propositions have there. Our goal, then, is to develop models that clearly reflect the fact that the change in truth value of a sentence at other worlds is due *not* to a change in their meaning but rather to a change in the world.

§2: Definition of the Philosophical Conception

To define these models, we extend the idea that propositions, modal or otherwise, are structured entities. Let us suppose that there is a fixed, denumerable set of expressible atomic propositions $\mathbf{P}^* = \{\mathbf{p}_1, \mathbf{p}_2, \ldots\}$. We form the set $\mathbf{P}$ of all the expressible propositions as follows, using the boldface symbols '$\boldsymbol{\varphi}$' and '$\boldsymbol{\psi}$' as variables ranging over propositions:[7]

[7] We are appealing here to a notion of 'expressibility' under which the set of expressible propositions is denumerable.

$\mathbf{P}$ is the smallest set satisfying the following conditions:

(a) Every element of $\mathbf{P}^*$ is an element of $\mathbf{P}$

(b) If $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathbf{P}$, so are $\mathbf{NEG}(\boldsymbol{\varphi})$, $\mathbf{COND}(\boldsymbol{\varphi}, \boldsymbol{\psi})$, and $\mathbf{NEC}(\boldsymbol{\varphi})$

We may suppose that **NEG**, **COND**, **NEC** are algebraic logical operations that harness simpler propositions into more complex ones. **NEG** and **NEC** are unary operations, whereas **COND** is a binary operation.

A *model* $\mathbf{M}$ *of* the domain $\mathbf{P}$ is any quadruple of the form $\langle \mathbf{W}, \mathbf{w}_\alpha, \mathbf{R}, \mathbf{ext} \rangle$, where $\mathbf{W}$ is a nonempty set of *possible worlds*, $\mathbf{w}_\alpha$ is a distinguished element of $\mathbf{W}$ called *the actual world*, $\mathbf{R}$ is a binary *accessibility* relation on $\mathbf{W}$, and $\mathbf{ext}$ is a binary *extension* function with domain $\mathbf{P} \times \mathbf{W}$ and range $\{\mathbf{T}, \mathbf{F}\}$ satisfying the following constraints (in which $\mathbf{ext}$ is indexed to its second argument):

(a) $\mathbf{ext}_\mathbf{w}(\mathbf{NEG}(\boldsymbol{\varphi})) = \mathbf{T}$ iff $\mathbf{ext}_\mathbf{w}(\boldsymbol{\varphi}) = \mathbf{F}$

(b) $\mathbf{ext}_\mathbf{w}(\mathbf{COND}(\boldsymbol{\varphi}, \boldsymbol{\psi})) = \mathbf{T}$ iff
either $\mathbf{ext}_\mathbf{w}(\boldsymbol{\varphi}) = \mathbf{F}$ or $\mathbf{ext}_\mathbf{w}(\boldsymbol{\psi}) = \mathbf{T}$

(c) $\mathbf{ext}_\mathbf{w}(\mathbf{NEC}(\boldsymbol{\varphi})) = \mathbf{T}$ iff
for every world $\mathbf{w}'$ such that $\mathbf{Rww}'$, $\mathbf{ext}_{\mathbf{w}'}(\boldsymbol{\varphi}) = \mathbf{T}$

Hereafter, we sometimes index the major elements of $\mathbf{M}$ as: $\mathbf{W_M}$, $\mathbf{R_M}$, and $\mathbf{M}\text{-}\mathbf{ext}_\mathbf{w}$, respectively. Notice that models are defined relative to the domain of propositions $\mathbf{P}$.

We now define notions of truth that apply to the propositions in $\mathbf{P}$. The notions $\boldsymbol{\varphi}$ is *true at* world $\mathbf{w}$ *in* $\mathbf{M}$ (in symbols: $\mathbf{M}, \mathbf{w} \models \boldsymbol{\varphi}$) and $\boldsymbol{\varphi}$ is *true in* model $\mathbf{M}$ (in symbols: $\mathbf{M} \models \boldsymbol{\varphi}$) are defined as follows:

$$\mathbf{M}, \mathbf{w} \models \boldsymbol{\varphi} =_{df} \mathbf{M}\text{-}\mathbf{ext}_\mathbf{w}(\boldsymbol{\varphi}) = \mathbf{T}$$

$$\mathbf{M} \models \boldsymbol{\varphi} =_{df} \mathbf{M}, \mathbf{w}_\alpha \models \boldsymbol{\varphi}$$

We define the notion $\boldsymbol{\varphi}$ is *valid* (in symbols: $\models \boldsymbol{\varphi}$) as follows:

$$\models \boldsymbol{\varphi} =_{df} \text{For every } \mathbf{M},\ \mathbf{M} \models \boldsymbol{\varphi}$$

Finally, we define, for set $\boldsymbol{\Gamma}$ of propositions, the idea that $\boldsymbol{\Gamma}$ *logically implies* $\boldsymbol{\varphi}$, as follows:

$\boldsymbol{\Gamma}$ *logically implies* $\boldsymbol{\varphi}$ $=_{df}$
For every $\mathbf{M}$, if $\mathbf{M} \models \boldsymbol{\psi}$, for each $\boldsymbol{\psi} \in \boldsymbol{\Gamma}$, then $\mathbf{M} \models \boldsymbol{\varphi}$

Now that truth, validity, and logical consequence have been defined for propositions, we may consider how language enters the picture. Let us take a standard language, $\mathcal{L}$, containing atomic formulas $p_1, p_2, \ldots$ and complex formulas of the form: $\neg\varphi$, $\varphi \rightarrow \psi$, and $\Box\varphi$. Notice here that we are using nonbold letters 'φ' and 'ψ' to range over sentences. An *interpretation* **I** of this language is simply any pair $\langle \mathbf{P}, \boldsymbol{\delta} \rangle$, consisting of: (a) the (fixed) domain **P** of expressible propositions defined above, and (b) a *denotation* function $\boldsymbol{\delta}$ that maps the formulas of $\mathcal{L}$ to propositions in **P** subject to the conditions that:

If $\varphi = \neg\psi$, $\boldsymbol{\delta}(\varphi) = \mathbf{NEG}(\boldsymbol{\delta}(\psi))$

If $\varphi = \psi \rightarrow \chi$, $\boldsymbol{\delta}(\varphi) = \mathbf{COND}(\boldsymbol{\delta}(\psi), \boldsymbol{\delta}(\chi))$

If $\varphi = \Box\psi$, $\boldsymbol{\delta}(\varphi) = \mathbf{NEC}(\boldsymbol{\delta}(\psi))$

It proves useful to single out an interpretation $\mathbf{I}^*$ that constitutes the *intended interpretation* of our language. $\mathbf{I}^*$ is simply the interpretation such that $\boldsymbol{\delta}_{\mathbf{I}^*}$ is defined on the atomic formulas of $\mathcal{L}$ as follows:

If $\varphi = p_i$, $\boldsymbol{\delta}_{\mathbf{I}^*}(\varphi) = \mathbf{p}_i$, for every i

Of course, other interpretations relax the definition of $\boldsymbol{\delta}_{\mathbf{I}}$ in a variety of ways: (a) $\boldsymbol{\delta}_{\mathbf{I}}$ need not map each atomic sentence to the correspondingly numbered atomic proposition, (b) $\boldsymbol{\delta}_{\mathbf{I}}$ need not map the atomic sentences *onto* $\mathbf{P}^*$, and (c) $\boldsymbol{\delta}_{\mathbf{I}}$ need not map the atomic sentences to atomic propositions.

We may now formulate the following general definitions of the meaningfulness of sentences φ and the expressibility of propositions $\boldsymbol{\varphi}$ relative to an interpretation **I**:

φ is *meaningful*$_{\mathbf{I}}$ $=_{df}$ $\exists \boldsymbol{\varphi} \in \mathbf{P}$ such that $\boldsymbol{\delta}_{\mathbf{I}}(\varphi) = \boldsymbol{\varphi}$

$\boldsymbol{\varphi}$ is *expressible*$_{\mathbf{I}}$ $=_{df}$ $\exists \varphi \in \mathcal{L}$ such that $\boldsymbol{\delta}_{\mathbf{I}}(\varphi) = \boldsymbol{\varphi}$

In other words:

A sentence φ is meaningful relative to interpretation **I** iff there is some proposition in **P** that φ denotes under **I**.

A proposition $\boldsymbol{\varphi}$ is expressible relative to interpretation **I** iff there is a sentence in $\mathcal{L}$ that denotes $\boldsymbol{\varphi}$ under **I**.

Of course, we have set things up so that in every interpretation **I**, all of the sentences of $\mathcal{L}$ are meaningful$_{\mathbf{I}}$. However, there are interpretations of the language in which some of the propositions in **P** are not expressible.

Now, it is straightforward to define related notions of truth at a world, truth, and validity, all of which apply to sentences. Such definitions presuppose a *fixed* interpretation of the language, for these subordinate notions of truth and logical truth are only properly defined for interpreted sentences. To emphasize that these notions are being defined relative to a fixed interpretation of the language, we use '$[\varphi]_{\mathbf{I}}$' to talk about the formula φ *under* the interpretation **I**. Though models **M** were originally defined relative to the domain **P**, it is now useful to suppose that they are defined relative to **I** itself, despite the fact that they have nothing to do with $\boldsymbol{\delta}_{\mathbf{I}}$. We shall not, however, index the models to **I**, since it is to be understood that they are defined for the fixed domain **P** in **I**. We then have the following definitions:

$\mathbf{M}, \mathbf{w} \models [\varphi]_{\mathbf{I}} =_{df} \mathbf{M}, \mathbf{w} \models \boldsymbol{\delta}_{\mathbf{I}}(\varphi)$

$\mathbf{M} \models [\varphi]_{\mathbf{I}} =_{df} \mathbf{M} \models \boldsymbol{\delta}_{\mathbf{I}}(\varphi)$

$\models [\varphi]_{\mathbf{I}} =_{df} \models \boldsymbol{\delta}_{\mathbf{I}}(\varphi)$

In general, a sentence under interpretation **I** will have a semantic property just in case the proposition it denotes under **I** has that property. We may also define a notion of logical consequence that holds between a set of interpreted sentences $\Gamma_{\mathbf{I}}$ (i.e., a set of sentences Γ each member of which is interpreted under **I**) and $[\varphi]_{\mathbf{I}}$ as follows:

$\Gamma_{\mathbf{I}}$ *logically implies* $[\varphi]_{\mathbf{I}}$ $=_{df}$
For every **M**, if $\mathbf{M} \models [\psi]_{\mathbf{I}}$, for every $\psi \in \Gamma$, then $\mathbf{M} \models [\varphi]_{\mathbf{I}}$

The idea that logic is primarily about propositions and secondarily about (interpreted) sentences can also be applied to the proof theory. Rules of inference can be defined as relations among propositions. Since propositions have a precisely understood structure, there should be no question as to when several propositions are related as the premises and conclusions of the classical rules. For example, the rule Modus Ponens is the rule that relates propositions $\boldsymbol{\varphi}$ and $\mathbf{COND}(\boldsymbol{\varphi}, \boldsymbol{\psi})$ to the proposition $\boldsymbol{\psi}$ as the premises and conclusion of the rule, respectively. A logic $\boldsymbol{\Sigma}$ can be defined as a set of propositions (containing zero or more propositions as 'axioms') which is closed under certain rules of inference. Theoremhood is a property of propositions: a proposition φ is a *theorem* of logic $\mathbf{\Sigma}$ (in symbols: $\vdash_\Sigma \varphi$) iff $\varphi \in \mathbf{\Sigma}$. We may define the consistency and completeness of logic $\mathbf{\Sigma}$ relative to the class of models $\mathbf{C}$ in the usual way:

$\mathbf{\Sigma}$ is *sound with respect to* $\mathbf{C} =_{df}$ If $\vdash_\Sigma \varphi$, then $\models_{\mathbf{C}} \varphi$

$\mathbf{\Sigma}$ is *complete with respect to* $\mathbf{C} =_{df}$ If $\models_{\mathbf{C}} \varphi$, then $\vdash_\Sigma \varphi$

And, finally, using our denotation function, we may develop the secondary notions of rule of inference, logic, and theoremhood, and derivability as they apply to *interpreted* sentences. This is straightforward, and we shall spend no further time developing the definitions.

Since we have, for simplicity, defined the domain of propositions $\mathbf{P}$ to be structurally identical to the set of formulas $\mathcal{L}$, there should be no question that all of the classical results of ML can be reconceived as results about propositions.

§3: Virtues of the Philosophical Conception

We now have a broader conception under which ML is concerned primarily with propositions and their truth conditions, and secondarily with interpreted sentences and their truth conditions. Validity and theoremhood are primarily notions that apply to propositions, and secondarily, to interpreted sentences. No matter what fixed interpretation $\mathbf{I}$ our language $\mathcal{L}$ has, the proposition a sentence denotes varies neither from model to model nor from world to world.

The reason that an atomic sentence 'p' has a different truth value in $\mathbf{w}_1$ and $\mathbf{w}_2$ is that the proposition it denotes changes truth values in these worlds, and this reflects a difference in the worlds themselves. Moreover, there is no room to suppose that the reason '$\Diamond\neg p$' is true is that there are worlds where 'p' denotes a (false) proposition different from the one it in fact denotes. So, the first virtue of the present conception is that it eliminates the undesirable interpretation of the fact that the truth value of a sentence can vary from one world to the next.

Thus, we may conceive of models $\mathbf{M}$ as models of the world (or rather, as models of the possibluum), rather than as models for a language. We can clearly separate the notion of an *interpretation of the language* from the notion of a *model of the world*. On the traditional conception, these

two notions are not clearly separated. And the reason they aren't is that the $\mathbf{V}$ function of traditional models does two jobs that should be separated. $\mathbf{V}$ not only interprets the language but also tells us what the facts are in the various possible worlds. It conflates the work of the denotation function $\boldsymbol{\delta}$ that forms part of an interpretation with that of the **ext** function that forms part of the model. $\mathbf{V}$ essentially *defines truth* for an uninterpreted sentence 'p' at the same time that it *interprets* 'p'. By separating out the $\boldsymbol{\delta}$ function and the **ext** function, we distinguish the interpretation of 'p' from the truth conditions of the proposition it denotes, and thereby clarify our picture of how language works. Truth can be conceived as a property of sentences only if the sentences are interpreted, i.e., are meaningful. And given our definitions in the previous section, we are able to say precisely wherein lies the meaningfulness of sentences.

Since we can distinguish interpretations from models, we can also distinguish the *intended* interpretation from the *intended* model *for* the intended interpretation.[8] If we suppose that $\mathbf{P}$ houses the propositions that are expressible in English and that each sentence φ of $\mathcal{L}$ abbreviates a sentence S of English, then we would define the intended interpretation $\mathbf{I}^*$ not as we did in the previous section, but rather as that interpretation in which $\boldsymbol{\delta}_{\mathbf{I}^*}(\varphi)$ is the proposition S in fact denotes. Consequently, the intended model $\mathbf{M}^*$ for the intended interpretation will be that model in which: (a) $\mathbf{W}$ contains all of the possible worlds that there are, in addition to the real world $\mathbf{w}_\alpha$, (b) $\mathbf{R}$ relates the members of $\mathbf{W}$ in just the way required by the real modal accessibility relation (whatever that turns out to be), and (c) **ext** distributes truth values to the propositions at each world in just the way required by the (modal) facts. Condition (c) requires, for example, that whenever S is a true English sentence, and φ abbreviates S in $\mathcal{L}$, then $\mathbf{M}^*\text{-}\mathbf{ext}_{\mathbf{w}_\alpha}(\boldsymbol{\delta}_{\mathbf{I}^*}(\varphi)) = \mathbf{T}$. Since $\mathbf{M}^*$ has to accomodate the true modal English sentences of the form "It is possible that S'" and "It is necessary that S'", $\mathbf{M}^*$ will be that model for the intended interpretation that correctly represents modal reality.

A second virtue of the present conception concerns the presence of a distinguished world in the models. This follows Kripke's original formulation, in which he includes such a distinguished element. By contrast, many recent developments of modal logic abandon this element. We shall not spend much time here discussing why we follow Kripke, since we have

[8]Compare Hanson and Hawthorne [1985].

described reasons for doing so at length in Zalta [1988b]. Note simply that without a distinguished actual world, one must define truth-in-**M** as truth at *every* world in **M** (as opposed to truth at the distinguished world of the model). But this seems to confuse truth with a kind of necessary truth. The lack of a distinguished world also affects the resulting definition of validity. If truth in **M** is defined as truth at all worlds in **M**, then validity (i.e., logical truth) must be defined as truth in every world in every model. In Zalta [1988b], we show that this definition is not equivalent to the notion of validity which Kripke originally defined and which we preserved in the previous section. We argued that there are valid propositions that are not necessarily true, and that this should make one extremely careful when moving from the model-theoretically defined claim that $\boldsymbol{\varphi}$ is valid to the metaphysical claim that the proposition $\mathbf{NEC}(\boldsymbol{\varphi})$ is true.[9] On the present model-theoretic analysis, the validity of a proposition depends on whether certain *nonmodal* facts about set-theoretic structures hold. Nothing follows about whether that proposition is metaphysically necessary, or true in all possible worlds, at least not unless certain other assumptions are made.

The main virtue of the philosophical conception of ML, however, is that it is more general. Our conception of ML *reduces* to the traditional one in the special case where the members of **P** are taken to be nothing other than the sentences of our formal language $\mathcal{L}$! But there is also a

[9]For those readers unfamiliar with my [1988b], here is an example of a logical truth that is not necessary. Consider the proposition *if it is actually the case that p, then p*, which we may represent using the actuality operator **A!** as: $\mathbf{COND}(\mathbf{A!p}, \mathbf{p})$. Suppose that for any model **M**, the extension of the proposition $\mathbf{A!}\boldsymbol{\varphi}$ (*it is actually the case that* φ) is the value **T** at a world **w** iff the extension of $\boldsymbol{\varphi}$ is **T** at the actual world. Notice that $\mathbf{COND}(\mathbf{A!p}, \mathbf{p})$ is logically true—in any model **M**, $\mathbf{COND}(\mathbf{A!p}, \mathbf{p})$ is true in **M**. But there are models in which the necessitation of this proposition is false. Consider the model $\mathbf{M}_1$ containing two worlds $\mathbf{w}_\alpha$ and $\mathbf{w}_1$ and such that:

(a) $\mathbf{Rw}_\alpha\mathbf{w}_1$

(b) $\mathbf{ext}_{\mathbf{w}_\alpha}(\mathbf{p}) = \mathbf{T}$

(c) $\mathbf{ext}_{\mathbf{w}_1}(\mathbf{p}) = \mathbf{F}$

Then, it is easy to prove that:

$$\mathbf{M}_1 \not\models \mathbf{NEC}(\mathbf{COND}(\mathbf{A!p}, \mathbf{p}))$$

To see this, just consider that in $\mathbf{M}_1$ there is a world accessible to the actual world in which the conditional $\mathbf{COND}(\mathbf{A!p}, \mathbf{p})$) is false, namely, $\mathbf{w}_1$. So the necessitation of $\mathbf{COND}(\mathbf{A!p}, \mathbf{p})$ is not true in $\mathbf{M}_1$. Thus, relative to $\mathbf{M}_1$, we have a case of a logical truth which is not necessary.

second way of reducing our conception to the old, namely, by requiring that our background theory of propositions identify necessarily equivalent propositions.[10] On the traditional account, such as that in Montague [1974], Lewis [1986], and elsewhere, **p** and $\mathbf{NEG}(\mathbf{NEG}(\mathbf{p}))$ must be the same proposition, for example, since they have the same truth value at each world. Moreover, since there is exactly one function that maps every world to the value **T**, the traditional account permits exactly one necessary proposition. By way of contrast, however, the fine-grained theory of propositions upon which our philosophical conception is grounded allows us to suppose that **p** and $\mathbf{NEG}(\mathbf{NEG}(\mathbf{p}))$ are distinct despite being necessarily equivalent. We may further suppose that there are an infinite number of distinct necessary truths. From this point of view, the traditional reconstruction of propositions as functions from worlds to truth values appears to confuse a proposition with a mathematical model of its behavior. Our more general approach separates propositions from their extensions, yet reduces to the Montagovian conception if one places a further constraint on the background theory of propositions, requiring that necessarily equivalent propositions be identical.

A final virtue of the philosophical conception is that it offers a more subtle understanding of logical equivalence. There are now *two* ways in which sentences φ and ψ can be logically equivalent. To distinguish these kinds of equivalence, let us digress momentarily to reflect upon the difference between formulas and terms. Formulas and terms are semantically significant in distinct ways. Formulas are expressions that have truth conditions. Formally speaking, they are the kind of expression for which *truth in a model* is defined. Terms, on the other hand, are expressions that have denotations. Formally speaking, they are expressions that fall in the domain of the function $\boldsymbol{\delta}$. Having truth conditions and having a denotation are different ways in which an expression can be semantically significant, and there is a separate notion of logical equivalence that is appropriate for each kind of semantic significance. Two formulas φ and ψ are logically equivalent just in case they logically imply each other (i.e., just in case φ and ψ are true in the same models). Two terms, however, are logically equivalent just in case they have the same denotation.

To make these definitions formally precise, we relativize these notions to an interpretation **I** of the language. By analogy with the notation '$[\varphi]_{\mathbf{I}}$',

[10]In other words, by adding the axiom $\Box(p \leftrightarrow q) \to p = q$ to our background theory of propositions.

we introduce the notation '$[\tau]_{\mathbf{I}}$' to refer to term τ *under* interpretation **I**:

Formulas $[\varphi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ are *logically equivalent* $=_{df}$
$[\varphi]_{\mathbf{I}}$ logically implies $[\psi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ logically implies $[\varphi]_{\mathbf{I}}$

Terms $[\tau]_{\mathbf{I}}$ and $[\tau']_{\mathbf{I}}$ are *logically equivalent* $=_{df}$ $\boldsymbol{\delta}_{\mathbf{I}}(\tau) = \boldsymbol{\delta}_{\mathbf{I}}(\tau')$

Thus, logical equivalence is defined in one way for interpreted formulas and in a second way for interpreted terms.

Recall now that the sentences of our language $\mathcal{L}$ are both formulas and terms. They are defined directly as formulas, and as such, receive truth conditions. But they also lie in the domain of the denotation function, and as such, constitute terms. Consequently, given an interpretation **I**, both of the above notions of logical equivalence apply to sentences. Henceforth, if two sentences are logically equivalent in the first sense, we shall say they are *formula-equivalent*. If they are logically equivalent in the second sense, we shall say they are *term-equivalent*.[11]

Whereas the notion of formula-equivalence has played the dominant role in the contemporary history of (modal) logic, term-equivalence (i.e., having the same denotation) is really the stronger kind of logical equivalence among sentences. It should be easy to see that if $[\varphi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ are term-equivalent, then they are formula-equivalent. Just fix an interpretation **I** and suppose that $[\varphi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ are term-equivalent. If $\boldsymbol{\delta}_{\mathbf{I}}(\varphi)$ just is $\boldsymbol{\delta}_{\mathbf{I}}(\psi)$, then any model in which $[\varphi]_{\mathbf{I}}$ is true is a model in which $[\psi]_{\mathbf{I}}$ is true, and vice versa. So $[\varphi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ logically imply each other, and so they are formula-equivalent. However, the formula-equivalence of two sentences, doesn't guarantee their term-equivalence. Again, fix an interpretation **I** and suppose $[\varphi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ are formula-equivalent. If $[\varphi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ are true in the same models, it doesn't follow that they have the same denotation under **I**. For $\boldsymbol{\delta}_{\mathbf{I}}(\varphi)$ and $\boldsymbol{\delta}_{\mathbf{I}}(\psi)$ may be distinct but logically equivalent propositions.[12]

[11] Strictly speaking, we should distinguish the formula φ from its counterpart term $[\lambda\,\varphi]$ (read: that-φ). In a language with enough expressive power to predicate properties of propositions, we should make such a distinction. The discussion that now follows should then be understood as remarks about the fact that the term-equivalence of $[\lambda\,\varphi]$ and $[\lambda\,\psi]$ in such a language would provide a stronger ground for the substitution principle than the formula-equivalence of φ and ψ.

[12] Examples of sentences that are formula-equivalent but not term-equivalent are easy to produce. Take the example we used a few paragraphs back. Propositions **p** and **NEG**(**NEG**(**p**)) are distinct propositions (this is something that is either a consequence of or consistent with any fine-grained theory of propositions). But, by the constraints on **ext**, they are true in the same models. Now pick any interpretation **I** and let it be a convention that $\boldsymbol{\delta}_{\mathbf{I}}(p) = \mathbf{p}$. It then follows by constraints on $\boldsymbol{\delta}$ that $\boldsymbol{\delta}_{\mathbf{I}}(p)$ and $\boldsymbol{\delta}_{\mathbf{I}}(\neg\neg p)$ logically imply each other. So $[p]_{\mathbf{I}}$ and $[\neg\neg p]_{\mathbf{I}}$ are formula-equivalent. But they are not term-equivalent, since $\boldsymbol{\delta}_{\mathbf{I}}(p)$ (i.e., **p**) and $\boldsymbol{\delta}_{\mathbf{I}}(\neg\neg p)$ (i.e., **NEG**(**NEG**(**p**))) are distinct.

The distinction between term-equivalence and formula-equivalence is especially useful for the analysis of other 'intensional' contexts. Consider the analysis of belief contexts, and take any two propositions that are intuitively distinct though logically equivalent; for example, (a) there is a barber who shaves all and only those who don't shave themselves, and (b) Clinton is president and it is not the case that Clinton is president. These two propositions are logically equivalent because in any model, they are both (necessarily) false. But someone could believe (a) and not (b), and so the fact that someone believes (a) does not imply that that person believes (b). In an intensional language and logic constructed in the spirit of our philosophical conception of ML, in which term- and formula- equivalence are distinguished, the formal representation of (a) would not imply the formal representation of (b). Here is why.

In an enriched formal language, sentence (a) might be denoted by the formula '$(\exists x)(Bx\ \&\ (\forall y)(Sxy \equiv\ \sim Syy))$' (let this be φ), and sentence (b) might be denoted by the formula '$Pc\ \&\ \sim Pc$' (let this be ψ). In any semantic treatment along the lines of the present conception, there will be interpretations **I** for which $[\varphi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ are formula-equivalent but not term-equivalent. This suggests an explanation of why a formal sentence such as $Bel(x, \varphi)$, asserting that person x stands in the belief relation to the proposition denoted by φ, would not imply the sentence $Bel(x, \psi)$, asserting that x stands in the belief relation to the proposition denoted by ψ. The further assumption that φ and ψ are term-equivalent is required for the valid inference from $Bel(x, \varphi)$ to $Bel(x, \psi)$. Belief contexts are sensitive to the identity of the proposition denoted by the embedded sentence. The formula-equivalence of two sentences (i.e., 'logically equivalent' in the traditional sense) is not a strong enough condition to preserve the identity of the propositions denoted, and so not a strong enough condition to guarantee the truth-preserving substitution of one sentence for another inside belief contexts. This approach may be contrasted with Montague's [1974] treatment of this case, in which $Bel(x, \hat{}\varphi)$ implies $Bel(x, \hat{}\psi)$, contrary to intuition.[13]

[13] The distinction between formula- and term-equivalence does not help to explain all of the cases of substitution failure that are problematic for Montague's system.

§4: A Consequence and Some Observations

An interesting consequence of our conception is that it reveals that the idea that 'modal contexts are intensional' to be somewhat confused. This idea has come to be one of the unquestioned truths of the analysis of language, but it may be largely in error. Here is a simple argument that shows that modal contexts are not intensional. Take the widely-accepted definition of intensionality found in van Benthem [1988]:

> Intensional Logic as understood here is a research program based upon the broad presupposition that so-called "intensional contexts" in natural language can be explained semantically by the idea of *multiple reference*. (p. 1)

So, in particular, if the propositional modal contexts of natural language are intensional, they are to be explained by the idea of a shift in denotation. But inspection of our definitions shows that the modal contexts of ML are not properly analyzed in such terms. If the modal contexts 'Necessarily,...' of natural language are analyzable in terms of formal sentences of the form $\Box\varphi$ in ML, and these formal sentences are not evaluated in terms of the multiple reference or shifting denotation of φ, it follows, by van Benthem's definition of intensionality, that such contexts of natural language are not intensional. On the philosophical conception of ML, the formal sentences $\Box\varphi$ and φ have a fixed denotation. The propositions these sentences denote do not vary, even when we evaluate the sentences at other worlds. The denotation function stands outside the model, and even if if one prefers to regard it as part of the interpretation/model complex, it is a unary function, not relativized to worlds. To use a well-known, vivid notion, the sentences of $\mathcal{L}$ are 'rigid designators', albeit in a somewhat vacuous sense. If we apply this analysis to the

It does not help to explain the 'hyperintensional' belief contexts; for example, why x can believe that Mark Twain is an author without believing that Samuel Clemens is an author. Here, the sentences 'Mark Twain is an author' and 'Samuel Clemens is an author' would seem to be term-equivalent. In Zalta [1983], (Chapter VI) and [1988a] (Chapters 9 – 12), there is an analysis of these cases in which the sentences in question are term-equivalent on the *de re* reading of the belief reports in which they are embedded, though not on the *de dicto* reading of such reports. In the *de dicto* readings, the sentences denote different singular propositions, which are distinguished by the fact that they have different abstract constituents (the distinct abstract objects that serve as the senses of 'Twain' and 'Clemens', respectively). Substitution of term-equivalent sentences inside *de dicto* belief contexts remains valid.

pure modal contexts of natural language, then it seems that the sentences embedded in these contexts should be regarded as 'rigid designators' as well.

Modal contexts even pass the classic pretheoretic test, failure of which identifies a context as intensional. On this classic test, a context of natural language is intensional if the substitution of 'coreferential terms' inside that context fails. But the modal contexts of ML pass the formal counterpart of this test. The term-equivalent sentences of ML are 'coreferential' terms, since such sentences constitute terms that denote the same proposition. Yet they may be validly substituted inside modal contexts. This is established by the following theorem:

> *Theorem*: If $\mathbf{M} \models [\Box\varphi]_{\mathbf{I}}$, and $[\varphi]_{\mathbf{I}}$ and $[\psi]_{\mathbf{I}}$ are term-equivalent, then $\mathbf{M} \models [\Box\psi]_{\mathbf{I}}$.

To simplify the proof, let us drop the subscripts that relativize everything to interpretation **I**. Thus, we drop the subscripts on formulas in the definitions of truth in a model and term-equivalence, as well as the subscript on the denotation function, it being understood that we shall be arguing with respect to a fixed interpretation **I**. Then:

> *Proof*: For an arbitrarily chosen **M**, suppose that $\mathbf{M} \models \Box\varphi$. Suppose also that φ and ψ are term-equivalent. We want to show: $\mathbf{M} \models \Box\psi$. By the definition of truth-in-**M** for sentences, we know that $\mathbf{M} \models \boldsymbol{\delta}(\Box\varphi)$. By the definition of $\boldsymbol{\delta}$, $\boldsymbol{\delta}(\Box\varphi) = \mathbf{NEC}(\boldsymbol{\delta}(\varphi))$. But from the fact that φ and ψ are term-equivalent, we know that $\boldsymbol{\delta}(\varphi) = \boldsymbol{\delta}(\psi)$. So $\mathbf{NEC}(\boldsymbol{\delta}(\varphi)) = \mathbf{NEC}(\boldsymbol{\delta}(\psi))$. But $\mathbf{NEC}(\boldsymbol{\delta}(\psi)) = \boldsymbol{\delta}(\Box\psi)$. So, by transitivity of identity, $\boldsymbol{\delta}(\Box\varphi) = \boldsymbol{\delta}(\Box\psi)$. Thus, $\mathbf{M} \models \boldsymbol{\delta}(\Box\psi)$, i.e., $\mathbf{M} \models \Box\psi$.

So substitution of term-equivalent sentences inside the modal contexts of ML preserves truth.[14] Thus, if the modal contexts of ML provide a correct analysis of the modal contexts of natural language, the latter contexts are not intensional.

So according to the classic definitions and tests for intensionality, modal contexts are not inherently intensional. Moreover, the fact that the modal contexts of ML are non-truth-functional does not entail that

[14] Of course, this is just a special case of a broader principle that allows us to substitute term-equivalent sentences in any context in ML.

they are intensional. The fact that modal contexts are non-truth-functional is established by the fact that $[\Box\varphi]_{\mathbf{I}}$ and $[\varphi \leftrightarrow \psi]_{\mathbf{I}}$ do not logically imply $[\Box\psi]_{\mathbf{I}}$ (it is straightforward to construct models that demonstrate this). But, again, this doesn't mean that modal contexts are intensional—they do not involve 'multiple reference' or 'shift in denotation'. So the present conception separates non-truth-functionality and intensionality.

Of course, there remains a sense in which modal *propositions* are *non*-extensional, namely, that $\mathbf{ext_w}(\mathbf{NEC}(\boldsymbol{\varphi}))$ doesn't depend on $\mathbf{ext_w}(\boldsymbol{\varphi})$. So, there is a *derivative* sense in which the modal contexts of language are nonextensional, namely, the truth of modal claim $\Box\varphi$ doesn't depend on the extension of the proposition denoted by φ. But it is important to distinguish this sense of 'nonextensional' from the classic tests used to identify a context as intensional. The present conception discriminates among grades of nonextensionality. The simple modal contexts of language exhibit nonextensionality without exhibiting intensionality of the sort defined by the classic definitions and tests.[15] This is a consequence of the more fine-grained analysis of language and modality.

[15]The quantified modal contexts of natural language that contain definite descriptions do attain the higher grade of nonextensionality—they are genuinely intensional. The famous puzzles about modal contexts involving such descriptions such as 'the morning star', 'the number of planets', and 'the teacher of Plato' (Quine [1961] and Kripke [1972]) are indeed cases where coreferential terms are not always intersubstitutable. In Zalta [1983] (pp. 99–106) and [1988a] (pp. 223–30), I treat such puzzles in a manner that is consistent with the present perspective. Though I agree with van Benthem that the best explanation of these contexts requires a shift in the denotation of the terms involved (i.e., I agree that they are genuine intensional contexts, according to his definition), I may disagree with him about the nature of the denotational shift. I do not suppose that the denotation of terms and sentences changes from world to world, but rather, since denotation is a unary notion not relativized to worlds, I postulate a shift in denotation *simpliciter*. In these genuine intensional contexts, definite descriptions receive an alternative denotation (an objectified individual concept), which play a role in a second interpretation of the sentence. Thus, I treat these descriptions, *not* as non-rigid designators that change their denotation from world to world, but rather as rigid designators which, inside certain modal contexts, shift from denoting one thing rigidly to denoting something else rigidly. In such modal contexts, they contribute something other than their ordinary denotation to the truth conditions of the sentence.

The belief contexts described in footnote 13 are also genuine cases of intensionality. Their analysis requires that we postulate a shift in denotation to account for the *de dicto* readings of the English sentences, for which substitution of coreferential names appears to fail.

Finally, there is a sense in which propositions, whether they are modal or not, are nonextensional, namely, that extensional equivalence (in the sense of either material or necessary equivalence) doesn't guarantee identity. Because of this, propositions are often referred to in the literature as 'intensional entities'. Consequently, a modal logic of propositions may be considered a logic of intensional entities, and to that extent, qualifies as an intensional logic. But then, ordinary *non*-modal propositional logic, if defined along the lines of the present philosophical conception, would constitute an intensional logic in that sense as well.[16]

I would like to conclude with two observations. The first concerns how the present conception of propositional modal logic differs from that developed in Bealer [1982], Menzel [1986] and [1990], and Zalta [1983] and [1988]. In these works, the authors construct languages and models for their intensional logics in which the domain of structured propositions forms a part of the models. In essence, we are given a language $\mathcal{L}$ and a model of the form $\langle \mathbf{P}, \mathbf{W}, \mathbf{w}_\alpha, \mathbf{R}, \mathbf{ext}, \boldsymbol{\delta} \rangle$; the domain of propositions does not form an independent structure, and instead of assigning sentences denotations in the separate domain of propositions, the denotation function $\boldsymbol{\delta}$ links the language directly into another part of the model. This still allows sentences to denote distinct, but equivalent, propositions, but notice that one may vary the model simply by varying either the domain of propositions or the denotation function.

However, when we try to develop the definitions of logical truth for *interpreted sentences* on this approach, a problem arises. To discuss interpreted sentences, we need to fix the domain of propositions $\mathbf{P}$ in the model and the denotation function that assigns every sentence a proposition. But for an interpreted sentence be logically true, we must be able to quantify over all models, since logical truth is truth in all models. The problem is that in addition to all the models that differ solely with respect to $\mathbf{W}$, $\mathbf{R}$, and $\mathbf{ext}$ function, there are all those models that have different domains $\mathbf{P}$ and different denotation functions $\boldsymbol{\delta}$. How then can we define logical truth for an interpreted sentence, i.e., for a sentence that requires us to fix the domain of propositions and the denotation function? The

[16]For those readers familiar with our background theory of propositions, we should mention that the identity of propositions p and q depends on the necessary equivalence of the *encoding* extension of the properties $[\lambda y\, p]$ and $[\lambda y\, q]$. That is, on our background theory of propositions, p and q are identical just in case $[\lambda y\, p]$ and $[\lambda y\, q]$ are necessarily encoded by the same objects. So there is still a sense in which extensionally equivalent propositions are identical. See footnote 5.

idea of an interpreted sentence requires us to hold part of the model fixed, but the definition of logical truth forces us to vary the models. Of course, a special and no doubt confusing indexing scheme could be developed to avoid this problem, I suppose, but the problem is avoided altogether under the present conception. By taking the domain of structured propositions and the denotation function out of the models, treating the former as a separate domain and the latter as the interpretation function, we can define interpreted sentences without fixing the models in any way.

The second and final observation is this. When it comes time to *apply* propositional modal logic for certain philosophical purposes, say, in the interpretation of modality in natural language or in the assertion of non-logical metaphysical principles, we may wish to pull other elements out of the models and add them as part of the fixed structure of propositions. For example, if for certain philosophical purposes we want to fix the size of $\mathbf{W}$, fix the properties of the accessibility relation, and explore non-logical principles that assume this modal structure, then it would be useful to take the domain of worlds, the actual world, and the accessibility relation out of the models and add additional structure to the domain of propositions. That is, for certain philosophical purposes, it would be useful to distinguish: the language $\mathcal{L}$, a structure $\langle \mathbf{P}, \mathbf{W}, \mathbf{w}_\alpha, \mathbf{R} \rangle$, and models $\mathbf{M}$ that consist solely of an **ext** function. In this way, models would vary only by the way each **ext** function maps out the space of possibilities. This may help capture the idea that it doesn't make sense to wonder whether there might have been more or fewer possible worlds than there in fact are.

Bibliography

Bealer, G., 1982, *Quality and Concept*, Oxford: Clarendon

Bealer, G., 1979, 'Theories of Properties, Relations, and Propositions', *The Journal of Philosophy*, **76**: 643-8

van Benthem, J., 1988, *A Manual of Intensional Logic*, Second Edition, Stanford University: Center for the Study of Language and Information Publications

Bull, R., and Segerberg, K., 1984, 'Basic Modal Logic', in *Handbook of Philosophical Logic: II*, D. Gabbay and F. Guenthner (eds.), Dordrecht: Reidel

Chellas, B., 1980, *Modal Logic: An Introduction*, New York: Cambridge

Dowty, D., Wall, F., and Peters, S., 1981, *An Introduction to Montague Semantics*, Boston: D. Reidel

Etchemendy, J., 1990, *The Concept of Logical Consequence*, Cambridge, MA: Harvard

Goldblatt, R., 1987, *Logics of Time and Computation*, Stanford: Center for the Study of Language and Information Publications

Hanson, W., and Hawthorne, J., 1985, 'Validity in Intensional Languages: A New Approach', *Notre Dame Journal of Formal Logic* **26**/1 (January): 9–35

Hughes, G. E. and Cresswell, M. J., 1968, *An Introduction to Modal Logic*, London: Methuen

Kripke, S., 1963, 'Semantical Considerations on Modal Logic', *Acta Philosophica Fennica*, **16**: 83–94

Kripke, S., 1972, *Naming and Necessity*, in *Semantics of Natural Language*, D. Davidson and G. Harman (eds.), Dordrecht: D. Reidel, 253–354; reprinted 1980, Cambridge, MA: Harvard

Lemmon, E. J. and Scott, D., 1977, *An Introduction to Modal Logic*, American Philosophical Quarterly Monograph Series, Number 11, Oxford: Basil Blackwell

Lewis, D., 1986, *On the Plurality of Worlds*, Oxford: Blackwell

McMichael, A., and Zalta, E., 1980, 'An Alternative Theory of Nonexistent Objects', *Journal of Philosophical Logic*, **9** (1980): 297–313

Menzel, C., 1990, 'Actualism, Ontological Commitment, and Possible World Semantics', *Synthese* **85**: 355–89

Menzel, C., 1986, 'A Complete, Type-free "Second-order" Logic and Its Philosophical Foundations', Technical Report No. CSLI–86–40, Stanford, CA: Center for the Study of Language and Information Publications

Montague, R., 1974, *Formal Philosophy: Selected Papers of Richard Montague*, Richmond Thomason (ed.), New Haven: Yale University Press

Myhill, J., 1963, 'An Alternative to the the Method of Extension and Intension', in *The Philosophy of Rudolf Carnap*, P. Schilpp (ed.), La Salle: Open Court, 299–310

Parsons, T., 1980, *Nonexistent Objects*, New Haven: Yale University Press

Quine, W., 1961, 'Reference and Modality', in *From a Logical Point of View*, New York: Harper and Row

Segerberg, K., 1971, *An Essay in Classical Modal Logic*, Philosophy Society and Department of Philosophy, University of Uppsala, Vol. 13

Zalta, E., 1993, 'Twenty-Five Basic Theorems in Situation and World Theory', *Journal of Philosophical Logic*, **22**: 385–428

———, 1988a, *Intensional Logic and the Metaphysics of Intentionality*, Cambridge, MA: MIT/Bradford

———, 1988b, 'Logical and Analytic Truths That Are Not Necessary', *The Journal of Philosophy* **85**/2 (February): 57–74

———, 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: D. Reidel