From a Modal Point of View

A Logical Investigation into Modalities De Dicto and De Re

By
Paul Harrenstein

October 8, 1998

Contents

Chapter 1

Introduction

There are basically two ways of doing modal logic. One way is by concentrating on a specific kind of modal expression, signifying in some sense a concept of possibility or necessity, and trying to make its formal characteristics explicit. In this manner one may discover that, when investigating, say, the notion of provability in Peano arithmetic as a modal concept, any formula of the form $\Box(\Box\varphi \to \varphi) \to \Box\varphi$ will be a theorem. This being quite a remarkable formula schema, it will not generally be true for other modal concepts. Other notions of modality may behave according to quite different formal principles. Eventually, one may come to realize that modal logic has many and widely divergent applications. This may in turn create the impression that different modal logics are only affiliated by a nebulous notion of a modal concept and the use of the symbols $\Box$ and $\Diamond$ in their respective formalizations.

Another way of doing modal logic is to start off with an abstract modal operator, $\Box$ say, and issue some formal principles governing its use and which are presumed to be common to any concrete modal concept. Necessitation (if φ is derivable so is $\Box\varphi$) and distributivity of $\Box$ over $\to$ ($\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$) are most often taken to be such principles. The idea is to develop a minimal modal logic, i.e. a logic such that any other logic formalizing a substantial modal concept is a conservative extension of it. Conceived in this manner, modal logic is a unifying program. A good example of this approach is Łukasiewicz [1953], who dismissed the characterization of modal logic as logical systems in which expressions such as 'necessary' and 'possible' occur, as being too vague. For similar reasons the propositional modal logic **K** — comprising the axiom scheme $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, the axioms for the truth functions, and *Modus Ponens* and Necessitation as the only rules of inference — is also of interest. The reason for this is not so much that it intuitively embodies such an enticing notion of modality but rather because it can be conceived of as a minimal modal logic. Widely divergent modal notions can be analysed in terms of the class of Kripke frames they characterize. The generality of **K** resides in the fact that it is both sound and complete with respect to the class of *all* Kripke frames.

The power of this second approach, however, should not be overestimated. It remains a pretty moot point whether modal concepts can be demarcated from non-modal ones by formal means only. It could even be queried whether $\Box$ when fully specified by the **K** axioms should be classified as *modal*. Moreover, the latter approach is, of course, not entirely independent of the former. The formal principles of $\Box$ should, after all, be shared by the concrete modal notions one wishes to analyse.

Adopting the second approach, one may come to wonder whether the principles governing the use of $\Box$ in a predicate logical language are identical to those of the propositional fragment, or whether $\Box$ also operates on the predicate logical structure of the formulae in its scope. The phenomena concerning the intensionality of modal contexts at a predicate logical level of analysis reveals that genuinely predicate logical principles of modality do exist.

The failure of substitutivity *salva veritate* of material equivalent sentences in modal contexts makes for the intensionality of propositional modal logics. This feature is inherited by most, if not all, modal predicate logics. Dealing with languages for modal predicate logic, however, the intensionality of modal contexts is more pervasive than this. Apart from substitutivity of material equivalent (closed) formulae, first-order predicate logic also grants substitution *salva veritate* of coreferential terms and substitution of co-extensive predicates. When modalities are added to the language none of these extensionality criteria are likely to be maintained. Most conspicuous is the failure of substitutivity of coreferential terms without loss of truth-value,[1] as can be illustrated by the following classic example that is due to Quine. Elementary arithmetic teaches us that:

(1.1) Nine is necessarily greater than five

whereas the astronomers persist in the truth of:

(1.2) The number of planets is nine.

However, circumstances are surely conceivable in which the number of planets happens not to be greater than five, thus falsifying:

(1.3) The number of planets is necessarily greater than five.

One of the philosophical puzzles connected with this failure of substitutivity of coreferential terms in modal contexts is to what extent examples like (1.1)–(1.3) constitute counterexamples against the much celebrated principle of indiscernibility of indenticals. The responses to this enigma have been diverse. Some logicians, like E.J. Lemmon, have maintained that (1.1)–(1.3) does indeed compromise the principle of the indiscernibility of identicals but that this is nothing to worry about. It is all part of growing up and doing intensional logic:

> [The failure of the principle of indiscernibility of identicals] will be unpalatable to many, but I believe it to be a paradox of intensionality that should be accepted on a par with the paradoxes of infinity that we have now come to accept (for example, that a totality may be equinumerous with a proper part of itself). ... The paradoxes of the infinite are paradoxical only because we normally think in terms of finite classes; this paradox of intensionality is paradoxical only because we normally think, with Leibniz, in extensional terms. (Lemmon [1963], p.98)[2]

[1] The failure of substitutivity of co-extensive predicates shall largely be ignored in this thesis.
[2] Cited in Cartwright [1971], p.140

Although this certainly speaks of the big, broad flexible outlook on Lemmon's part, I, with Quine, tend to think of the principle of the indiscernibility of identicals as one of the fundamental principles governing identity that should not be dismissed so easily.

Quite another response would be to point out that there is some way in which the inference from (1.1) and (1.2) to (1.3) is valid. If the integer nine can be attributed the property of being necessarily greater than five, so can the actual number of planets, it being identical to nine. On this reading (1.3) is true rather than false, let alone that (1.1)–(1.3) constitute a counterexample against the principle of indiscernibility of identicals. The apparent contingency of (1.3), however, can also be made sense of, if it is understood as expressing that of necessity there cannot be fewer than six planets. Construed this way, (1.3) does not attribute necessary properties to the object 'the number of planets' refers to, *viz.* the integer nine. Rather, it conveys that the sentence (1.4) is true in all possible circumstances:

(1.4) The number of planets is greater than five.

Again there is no infringement of the principle of indiscernibility of identicals. This time, however, because the object to which necessity is attributed in (1.3) is not the same object as the one that is said to be identical to nine in (1.2).

The classic difference between modality *de re* and modality *de dicto* can thus be invoked to avert the threatening break-down of the principle of indiscernibility of identicals. This distinction is traditionally between a necessary (or possible) property being attributed to an object (the *res*) and necessary (or possible) truth being attributed to a statement or sentence (the *dictum*).

Indulging in talk about possible worlds as a heuristic device, the distinction could be formed a mental picture of as follows. When evaluating a *de re* statement first the references of the referring terms in our world are determined. Only then it will be verified whether those objects possess the properties attributed to them in the various possible worlds. The statement can then be taken to concern the object denoted by the referring term in the actual world. In the case of a *de dicto* statement, on the other hand, the various references of the referring terms in the various accessible worlds are relevant. In fact, the evaluation of a *de dicto* modal statement proceeds by investigating in which of the possible worlds the sentence in the scope of the modal operator is true. Having put it in this way, it can be appreciated that the *de re*-*de dicto* distinction concerns the interpretation of referring expressions. If so, it would seem that a statements could contain several terms, some of which should interpreted *de re* and others *de dicto*. An example may be useful to illustrate this point.

(1.5) The President of the United States is necessarily the President of the United States.

(1.5) can mean several things. First, it could be taken to be *de dicto* for both the first and second occurrence of 'the President of the United States'. In that case it is most certainly true, since:

(1.6) The President of the United States is the President of the United States

is true in all possible worlds irrespective of who is the President of the United

States in those possible worlds. If both occurrences of 'the President of the United States' in (1.5) are taken *de re*, then again (1.5) is true, be it now because no object can fail to be self-identical, and the actual President of the United States is no exception. If American elections are fair to some degree, however, any candidate could have lost the elections, and so could the actual President. If the present President of the United States had lost the elections, not *he* but his rival would have been President of the United States. In that case the (actual) President of the United States would not have occupied the office of President of the United States. Understood this way, i.e. if the first occurrence of 'the President of the United States' is taken *de re* and the second *de dicto*, (1.5) is false. It is in this sense that the *de re-de dicto* distinction applies to the interpretation of referring terms.

Having distinguished between modality *de dicto* and modality *de re*, one should be cautious not to divorce these notions to too great an extent. The distinction between modality *de dicto* and modality *de re* should not be conceived of as being on a par with, e.g., the difference between physical and mathematical necessity. The contrast between these latter modalities seems somehow to reside in the different possibilities they acknowledge. Some state of affairs might very well be mathematically possible and yet be reckoned outright absurd by all physical standards. The difference between *de re* and *de dicto*, however, does not seem to relate to the possibilities that should or should not be recognized. Rather, it concerns the way we speak *about* what is acknowledged to be possible.

The distinction between *de re* and *de dicto* readings has frequently been explained in terms of the relative scope of quantifiers and modal operators.[3] The difference of interpretation of *de re* and *de dicto* readings of a modal sentence is then conceived of as spin-off to the interaction between quantifiers and modal operators. The sensitivity of the modal operators for the predicate logical structure of the formulae in their scope is taken care of by the quantifiers. In this respect it may appear that developing a logical system that combines both modality and quantification is a most promising course to take. However manageable any such scheme might appear, its execution will be fraught with peril. Simply lumping the axioms of quantification theory and propositional modal logic (of whatever preferred system) together does not correspond to an intuitively acceptable semantics. For reasons that will presently become clear, any such approach, moreover, may even jeopardize its own intelligibility by annihilating modal distinctions. If one tries to resolve matters from a semantical angle, one will readily be confronted with a plethora of choices as to the interpretation of the quantifiers, the identity sign, and referring terms. The point is that how these choices should be made cannot be decided in advance by taking recourse to the principles of the underlying logical systems. Moreover, these choices can be made largely independently. Consequently, a 'bewilderingly' large number of quantified modal logics have been proposed, none of which have attained a canonical status.[4]

[3] Cf. Smullyan [1948] in case the referring expressions can be analysed in terms of definite descriptions. See also chapter 2 of this thesis. If the referring expressions cannot be thus analysed, e.g. in the case of Kripkean proper names, it would seem that variables and the identity sign could be employed in order to achieve the distinction between *de re* and *de dicto* readings by means of quantifiers. I am indebted to Jelle Gerbrandy for pointing this out to me.

[4] For an overview of the quantified modal logics that have been advanced, the reader be referred

One of the ideas underlying this master's thesis is that the *de re-de dicto* distinction should be taken at face value. The possible non-equivalence of *de re-de dicto* readings of modal statements is taken to be a fundamental feature of any modal language with a predicate logical structure. Any modal predicate logic should comprise principles governing the use of $\Box$ that relate to this distinction. The question is, of course, which principles these are. That one could wish his modal predicate logic to be a conservative extension of a propositional modal logic goes almost without saying. This, however, is not quite so much as saying that there are no principles governing the use of modalities that concern predicate logical structure apart from their interaction with the quantifiers. In this thesis I will endeavour to cope with these predicate logical principles of modalities without taking recourse to quantification theory.

Speaking informally about formalized languages, the distinction between *de re* and *de dicto* attribution can, in fact, be extrapolated as to apply to any predicate logical language containing both a sentential operator $\otimes$ and λ-operators. The distinction can then be represented syntactically as follows:

(1.7) $(\lambda x. \otimes \varphi)(t)$ (*de re*)

(1.8) $\otimes((\lambda x.\varphi)(t))$ (*de dicto*)

The λ-operator is here conceived of as a device for forming predicates out of formulae. The extension of a λ-term $\lambda x.\varphi$ is assumed to be the class of values for x that satisfy φ. (1.7) exemplifies *de re* attribution in the sense that the property denoted by $\lambda x.\varphi$ is attributed to the object denoted by t, whereas in (1.8) $\otimes$ operates on the formula $(\lambda x.\varphi)(t)$ as a whole. The latter could be taken as attribution *de dicto*. Being equipped with any such language, one could formulate the principle of the indiscernibility of identicals as follows:

For all λ-terms $\lambda x.\varphi : t \equiv t' \rightarrow ((\lambda x.\varphi)(t) \rightarrow (\lambda x.\varphi)(t'))$

Given λ-conversion and -abstraction — $(\lambda x.\varphi)(t) \leftrightarrow [t/x]\varphi$ if t is free for x in φ, for some suitable notions of substitution and of a term being free for a variable — substitutivity of coreferential terms *salva veritate* can be seen to follow immediately from the indiscernibility of indenticals; be it only if the respective terms are free for substitution:

$t \equiv t' \rightarrow ([t/x]\varphi \rightarrow [t'/x]\varphi)$ $\qquad t, t'$ free for x in φ

λ-operators are, of course, referential contrivances on a par with the quantifiers. As such, an analysis of the *de re-de dicto* distinction in terms of λ's cannot be expected to avoid the problems of quantified modal logic. Yet, by presenting matters in this way, one can readily come to appreciate the relation that obtains between substitutivity and the *de re-de dicto* distinction. If the sentential operator $\otimes$ is extensional, e.g. if $\otimes$ betokens truth-functional negation, λ-conversion can be applied to prove the equivalence of (1.8) and (1.7). As has been stressed above, however, if $\otimes$ represents a modal notion, (1.7) and (1.8) will not in general be logically equivalent, let alone that λ-conversion can be invoked to prove this.

There are two ways how λ-conversion can fail to identify (1.7) and (1.8). Firstly, the λ-operator λx may fail to bind a free x in φ. If so, $\lambda x. \otimes \varphi$ will either be the universal property, holding of all objects, if $\otimes\varphi$ is true, or the absurd property, holding of no objects whatsoever, if $\otimes\varphi$ is false. In either case $\otimes$ can be said to

to Garson [1984]

create a *referentially* opaque context. A rather more interesting possibility is that, although λx does bind an occurrence of x in φ, and $\otimes$ is referentially transparent as such, λ-conversion fails to result in a formula equivalent to (1.8) because t is not free for x in $\otimes\varphi$. If so, although reference succeeded, substitution failed. $\otimes$ can then be said to create a *substitutionally* opaque context.[5] Any logical theory dealing with intensional operators should in one way or other be able to account for the difference between substitutional or referential opacity.

If now a modal context, $\Box$, is both fully substitutionally transparent, i. e. if each term can be substituted for a coreferential one in the scope of $\Box$, the intelligibility of quantified modal logic is in jeopardy. In *Word and Object* (pp.197-8) Quine put forward the following argument, which is meant to show that in any modal logic that allows for both substitutivity *salva veritate* and quantifying into modal contexts modal distinctions collapse. $\varphi \to \Box\varphi$ will be derivable in any such modal logic, he argued, and since in most interesting modal logics $\Box\varphi \to \varphi$ is an axiom, so will $\varphi \leftrightarrow \Box\varphi$. Modal distinctions would thus be annihilated. To appreciate the alledged theoremhood of $\varphi \to \Box\varphi$ consider:

$$t \equiv t' \to ([t/x]\Box t \equiv x \to [t'/x]\Box t \equiv x) \tag{1.9}$$

which for each pair of terms t and t' is a mere instance of the principle of substitution *salva veritate*. Since $\Box$ is assumed to be completely transparent in both substitutional and referential respect, we may conclude:

$$t \equiv t' \to (\Box t \equiv t \to \Box t \equiv t'). \tag{1.10}$$

Furthermore, since $\Box t \equiv t$ is a theorem of any modal logic, it can be left out, sanctioning the conclusion that:

$$t \equiv t' \to \Box t \equiv t' \tag{1.11}$$

Now take for t and t' respectively $\imath x \cdot x \equiv a$ and $\imath x \cdot (x \equiv a \wedge \varphi)$ for some φ containing no free x.[6] $(\imath x \cdot x \equiv a) \equiv (\imath x \cdot (x \equiv a \wedge \varphi))$ and φ are logically equivalent:

$$(\imath x \cdot x \equiv a) \equiv (\imath x \cdot (x \equiv a \wedge \varphi)) \leftrightarrow \varphi. \tag{1.12}$$

If φ is true, it follows that $(\imath x \cdot x \equiv a) \equiv (\imath x \cdot (x \equiv a \wedge \varphi))$. By (1.11), moreover, we know that this will be necessarily so, i. e.:

$$\varphi \to \Box((\imath x \cdot x \equiv a) \equiv (\imath x \cdot (x \equiv a \wedge \varphi))). \tag{1.13}$$

Finally, by replacement of logically equivalent formulae and (1.12) the disastrous:

$$\varphi \to \Box\varphi \tag{1.14}$$

may be obtained.

In view of this argument it might seem mysterious that logicians have succeeded in advancing formal semantical systems for quantified modal logics at all. One way is by meeting Quine's charge with stout denial: modal contexts simply cannot be substitutionally transparent with respect to *all* terms. Føllesdal[7] has pointed out that the disastrous consequences of Quine's argument can be averted if one distinguishes carefully between substitutional opacity and referential opac-

[5] Føllesdal makes a roughly analogous distinction between referential transparency and extensional opacity.

[6] If the use of the $\imath$-operator disturbs some readers, rest assured that the argument could also, though less intelligibly, be given using function symbols with fixed interpretations.

[7] Cf. Føllesdal [1969] and Føllesdal [1986].

ity. Although referential opacity entails failure of (non-vacuous) substitution *salva veritate*, the converse does not hold. Accordingly, one can distinguish between singular terms such as variables and individual constants, for which a modal operator is presumed to be transparent, and general terms, such as, e. g. functional terms and those created by means of the $\imath$-operator, for which modal contexts remain opaque. Modal operators are then referentially transparent in virtue of the singular terms and substitutionally opaque because general terms cannot in general be substituted. Differentiating singular terms, such as variables and individual constants, on the one hand and general terms on the other, is the way out of the predicament Quine puts the modal logician in:

> [I]t seems that what one has to do is to accept the thesis and still find a way of avoiding Quine's calamitous conclusion, that is one must find a semantics for modal contexts to which these contexts are at the same time *referentially transparent* (as required for quantification) and *extensionally opaque* [i.e. in my terminology *substitutionally opaque*]; that is, co-extensional expressions must not be interchangeable (since such interchangeability would amount to precisely to the collapse of modal distinctions warned against by Quine).
>
> From this it is immediately clear that a satisfactory semantics for the modalities must distinguish between expressions which refer (singular terms) and expressions which have extension (general terms and sentences, the extension of a sentence being its truth-value). (Føllesdal[1969], p.179)

The inference to $\varphi \rightarrow \Box\varphi$ is then impaired because by taking t and t' representing respectively the general terms $\imath x \cdot x \equiv a$ and $\imath x \cdot (x \equiv a \wedge \varphi)$, one cannot derive (1.10) from (1.9) any longer.

Against thus aligning the distinction between substitutional transparency and opacity with a division in the set of terms, can be brought in that it comes down to a mere avoidance of the problem. It fails to explain why modal contexts should be transparent for referential terms and not for general terms apart from merely avoiding the cataclysmic conclusions of Quine's argument.

An obvious repercussion of this approach is, moreover, that unbound variables and individual constants always get a *de re* reading in intensional contexts whereas general terms will have to be understood *de dicto* when occurring in an intensional context. It may be argued that this is not really something to be seriously concerned about. The semantical behaviour of proper names, demonstratives, pronouns and the like in natural languages may be pointed out to correspond to that of Føllesdal's referring terms in many important respects. The former may accordingly be taken to be the natural language analogues of the latter. Kripke's and Kaplan's papers on names and demonstratives may be invoked in support of this view. Notwithstanding the undeniable merit of Kripke's and Kaplan's philosophical exertions, this argument makes the whole enterprise of modal predicate logic dependent upon a specific interpretation of the formal framework and on the tenability of particular philosophical views. From a formal point of view it is not much more than a philosophical fig leaf.

The fundamental idea of this master thesis is to conceive of modal operators as sentence operators that bind terms in their scope in much the same manner as quantifiers can bind variables (chapter 2). Substitution of a term t for a coreferential can then be blocked in some cases because t occurs in the scope of a modal operator that binds it. This sensitivity of the modal operators to the predicate logical structure of the formulae in their scope will be represented syntactically by indexing them with the terms they do not bind. Modal contexts are assumed to be both referentially transparent for the terms the modal operator does not bind, and referentially opaque for the others. Whether one term can be replaced by another in a modal context depends on whether both terms occur in the index. Being not partial as to which terms are allowed to occur in the index of a modal operator, no gulf is created between the alleged classes of referential and general terms. Still, subsequently providing a suitable semantics for these indexed modal operators (chapter 3) will reveal that, modal distinctions do not collapse, not even if a modal operator is permitted to be indexed by the set of all terms, thus yielding both substitutional and referential transparency. For the interpretation of these indexed modal operators, the intuitive *de re-de dicto* distinction will serve as a starting point. The resulting semantics for *de re* modality proves to bear some important resemblances to Kaplan's views on directly reference (chapter 5). Moreover, the quantifiers themselves can be provided a modal foundation by pointing out how they could be conceived of as specific kind of modal operators (chapter 6).

Chapter 2

Philosophical Background

2.1 Substitution *salva veritate*

About all existing things it seems at least two things can be said without indulging in all too exotic metaphysics. The first, which is also known as Butler's Dictum, conveys that everything is identical to itself and to no other thing.[1] The second says that everything which has a certain property shares that property with everything to which it is identical. The latter is, of course, the celebrated principle of the indiscernibility of identicals.

On a more semantical level it would look as if these metaphysical views can be represented by Leibniz' Law of the substitutivity of coreferential terms *salva veritate*. This principle states that whenever an identity statement involving two terms is true, these terms may be substituted for each other in any statement without changing its truth-value. The way this semantical thesis is related to the metaphysical views espoused above could be put as follows. If two terms refer to the same object, this object remains identical to itself and to itself only, irrespective of the term used to refer to that object. Therefore, if the object referred to by a certain term can be said to possess a particular property, the same property can be attributed to the object referred to by a coreferential term, the objects denoted by both terms being identical. This semantical intuition has already been voiced by Shakespeare through the mouth of Juliet: "What's in a name? that which we call a rose By any other name would smell as sweet" (Shakespeare: *Romeo and Juliet*, II,ii). Thus it would seem that if a statement is used to say something about an object referred to in that statement by a certain expression, the same thing could just as well be said about that very object deploying the statement that results from substituting the expression in question by a coreferential one.

Whereas the metaphysical views have the appearance of being well-nigh inca-

[1] Not everybody agrees that it sensibly can be *said*, though. Compare:

Beiläufig gesprochen, von *zwei* Dingen zu sagen, sie seien identisch, ist ein Unsinn, und von *einem* zu sagen, es sei identisch mit sich selbst, sagt gar nichts.	Roughly speaking, to say of *two* things that they are identical is nonsense, and to say of *one* thing that it is identical with itself is to say nothing at all. (Wittgenstein [1922],5.5303)

pable of arousing serious philosophical controversy, Leibniz' Law has frequently been looked askance at. One should observe that not any notion of substitutivity serves as well as another with respect to linking up the metaphysical principle of the indiscernibility of identicals and the semantical claim of substitution *salva veritate*. Mere replacement of one string of symbols for another will not do. Any such notion will not be able differentiate *vulgar* and *accidental* occurrences of expression, a distinction that derives from Kaplan:

> Expressions are used in a variety of ways. Two radically different ways in which the expression 'nine' can occur are illustrated by the paradigms:
>
> (2.1) Nine is greater than five
>
> (2.2) Canines are larger than felines.
>
> Let us call the kind of occurrence illustrated in (2.1) a *vulgar* occurrence, and that in (2.2) an *accidental* occurrence (or following Quine, an orthographic accident). For present purposes we need not try to define either of these notions; but presumably there are no serious logical or semantical problems connected with occurrences of either kind. The first denotes, is open to substitution and existential generalization, and contributes tot the meaning of the sentence which contains it. To the second, all such concerns are inappropriate. (Kaplan[1969],p.206)

Defining a notion of substitution in such a way that it does not allow for substituting an expression for another if the former has an accidental occurrence would not seem, however, a particularly complicated task from a logical point of view.[2]

The semantic behaviour of an expression between quotation marks is comparable to that of an expression having an accidental occurrence. Quotation is a device for forming a name of an expression by putting that expression between inverted commas. Thus 'Cicero' is a name that denotes the great Roman orator, whereas the result of putting that name between quotation marks once again, using double quotes this time, "'Cicero'"refers to a *name* of Cicero, *viz.* 'Cicero' and not to man himself in particular. For 'Tully' and "'Tully'" much the same applies. One should carefully distinguish between use and mention here. This accounts for the fact that, although Cicero and Tully happened to be one and the same person, 'Cicero' cannot be replaced by 'Tully' in (2.3) without it changing its truth value:

(2.3) 'Cicero' contains six letters.

(2.3) says about the name 'Cicero' that it contains six letters; something that is not true of 'Tully', which is an entirely different name. Replacing 'Cicero' by 'Tully' in (2.3) would coincide with substituting "'Tully'" for "'Cicero'", which would by no means amount to a substitution of coreferential expressions. So any suitable notion of substitution should allow expressions to be replaced only if they are used and not if they are merely mentioned.[3]

---

[2] I think this distinction between accidental and vulgar occurrences can also be invoked in cases of homonymy. If A and B are homonyms, A can be said to occur accidentally in B, though B, of course, cannot. In this way homonymy can be considered a borderline case of accidental occurrence. This may also account for many arguments that hinge on homonymy having an air of triviality or irrelevance from a semantical point of view.

[3] For my use of 'use' and 'mention', the reader be referred to Quine [1940], §4.

Similar concerns are apposite with respect to cases in which essential use is made of the fact that an object cannot be mentioned without using an expression referring to that object, such as in:

(2.4) Giorgione was so-called because of his size,

(2.5) Cæsar is referred to by the name 'Cæsar' in this sentence.

In spite of the truth of respectively:

(2.6) Giorgione is Barbarelli,

(2.7) Cæsar is the author of *De Bello Gallico*,

(2.8) and (2.9) are false:

(2.8) Barbarelli is so-called because of his size.

(2.9) The author of *De Bello Gallico* is referred to by the name 'Cæsar' in this sentence.

Yet, the confused way in which expressions are both used and mentioned in these cases can easily be disentangled, as witness the equivalence of (2.4) and (2.5) to respectively the unproblematic (2.10) and (2.11):

(2.10) Giorgione was called 'Giorgione' because of his size,

(2.11) Cæsar is referred to by the name 'Cæsar' in (2.5).

The occurrences of 'Giorgione' and 'Cæsar' which do not occur between quotation marks *can* be substituted for respectively 'Barbarelli' and 'the author of *De Bello Gallico*' in (2.10) and (2.11) without loss of truth value. Moreover, the threatening breakdown of the principle of indiscernibility of identicals can thus be averted. Neither 'to be so-called' nor 'referred to by the name "Cæsar" in this sentence' should be conceived of as one-place predicates attributable to objects; rather do they denote two-place relations holding between objects and their names and, respectively, objects and sentences in which names of that object occur.

More interesting cases in which Leibniz' Law (apparently) fails are those involving modalities and propositional attitudes. Drawing again from a stock of well-worn examples, consider:

(2.12) Necessarily, nine is greater than five

(2.13) 'Nine is greater than five' is a truth of arithmetic

(2.14) Hegel believed nine to be greater than five

All of these are true or could very well have been. Yet, substituting 'nine' by a co-referential expression, say 'the number of planets', turns them into conspicuous falsehoods:

(2.15) Necessarily, the number of planets is greater than five

(2.16) 'The number of planets is greater than five' is a truth of arithmetic

(2.17) Hegel believed the number of planets to be greater than five

In spite of this failure of substitution *salva veritate* the occurrence of 'nine' in the above examples is not so blatantly accidental as in (2.2). In an important sense, 'nine' makes a significantly more substantial contribution to the meaning of the sentences (2.12)–(2.14) then 'Cicero' does in (2.3) or 'can' in (2.2). Adopting Kaplan's terminology, 'nine' is said to have an *intermediate* occurrence in (2.12)–(2.14).

The question is how the failure of substitution *salva veritate* in intensional contexts should be properly understood. The cases of the modalities and the propositional attitudes are not completely analogous. The propositional attitudes, most particularly those involving belief, bring their own problems, which are not entirely germane to the modalities. Here the modal case will be focussed on. But even restricted as such, the issue has evoked many and widely divergent reactions from both philosophers and logicians, such as Quine, Frege and Arthur Smullyan.

Quine took examples like (2.12)–(2.17) as indicating that 'nine' and the 'the number of planets' do not occur *purely referentially*. It indicates that it is not just the references of the denoting expressions in extensional contexts that are relevant to the truth value of the sentence. Somehow, the intensional contexts are responsible for that; they are referentially opaque.

In 'Three Grades of Modal Involvement' (Quine [1953c]) Quine distinguishes three ways in which one could represent modal expressions in one's logic or semantics. At the first level of logical analysis modal expressions are conceived of as semantical predicates that can be attributed to names of statements. 'Necessary' and 'possibly' are thus taken to be ellipsis for 'it is necessary that', respectively 'it is possible that'. Quine construed 'that' as "an operator that attaches to a sentence to produce a name of a proposition".[4] Substitution within a modal context is then just as spurious as substitution of expressions occurring between quotation marks. As such, quotation is generally recognized as the opaque context *par excellence*. Under this conception, quantifiers cannot bind variables within names of statements. If x is a variable occurring in a statement $\varphi(x)$, then in its name, '$\varphi(x)$', the letter x is mentioned rather than used.[5] It would be as if we were to derive (2.18) from (2.3) by existential generalization:

(2.18) $\exists x$ 'x' contains six letters.

"'x' contains six letters" merely conveys so much as that the 24th letter of the alphabet contains six letters, which it does not.

Alternatively, one could conceive of 'necessary' and 'possibly' as statement operators, depicted by $\Box$ and operating on closed sentences. This might not seem to amount to a substantial departure from the previous case, it being apparently just a notational variant in which the quotation marks are suppressed. The opacity troubles keep lingering, be it that they had better be compared with the case where 'Barbarelli' and 'the author of *De Bello Gallico*' are substituted for 'Giorgione' and 'Cæsar' in respectively (2.4) and (2.5). The syntactic structure may suggest that the referential expressions within the scope of $\Box$ occur purely referentially. Examples like (2.12) and (2.15), however, show that this is not in general the case.

This second grade of modal involvement, however, may tempt one to lift the referential veil somewhat more by analysing the modalities as sentence operators that can also be attached to open sentences. If so, one had better proceed with care. The argument on page 6, above, indicates how overexposure might prove fatal. But even if one takes these counsels to heart, one can be certain that in doing so one would only unmask Aristotelian Essentialism in one of its guises. If

[4]Cf. Davidson and Hintikka [1969], p.344. I suppose that by 'proposition' Quine here means a bearer of truth rather the kind of elusive entity that sentences are sometimes alleged to express and the feasibility of which Quine so emphatically denied.

[5]I have here taken a liberty in quoting the schematic letter φ.

one were to derive:

(2.19) $\exists x$ necessarily, x is greater than five

from (2.12) by an application of existential generalization, in virtue of which object would this inference be licensed? Seemingly, any object identical with nine but somehow different from the number of planets would do. Since, however, the number of planets happens to be identical to nine, the prospects for specifying any such object are obviously rather bleak, to say the least. Alternatively, one could try to make sense of (2.19) by pointing out that 'nine' is somehow a better specification of the number nine in intensional contexts than 'the number of planets' is. But even if this feat could be achieved and if the truth of (2.19) is granted, it would seem that one is committed to Aristotelian essentialism in the sense that one can come to conceive of necessity as somehow residing in objects, rather than in language. Quine, not willing to subscribe that view by any means, concluded that quantified modal logic was ill-conceived and philosophically unsound.

Adherents of a second view, who derive their arguments and methods from Frege's writings, deny that expressions in intensional contexts have got their usual denotation. Rather, the denotations of expressions in intensional contexts are shifted, their reference is oblique. In this sense modal contexts would be comparable to quotation. Quotation could alternatively be looked upon, not so much as a name forming device, but rather as an operation on expressions that shifts their reference to one of their names, e. g. the reference of 'Cicero' would be shifted from Cicero, the great Catiline denouncer, to 'Cicero', a name of his. Operators creating intensional contexts would function in much the same way, be it that they would do not so much as shift the reference of expressions in its scope to names of those expressions but rather they would shift the reference of an expression from its usual denotation to its usual sense, by means of which its reference can be determined. If one contemplates substituting a co-referential expression for 'nine' in (2.12)–(2.14), one should bear in mind that 'nine' here denotes its sense. Accordingly, any eligible substitute for 'nine' would also have to refer to the usual intension of 'nine', i. e. the intension of 'nine' in extensional contexts, when occurring within intensional contexts like those of (2.12)–(2.14). 'The number of planets', differing widely in meaning from 'nine' and only incidentally having the same denotation as the latter, will not qualify in this respect. Consequently, (2.12)–(2.17), should not be regarded as genuine cases of substitution of *coreferential* expressions, and as such fail to refute Leibniz' Law.

On this account, quantification into modal contexts does make sense, be it that the variables should then range over some kind of intensional object, i. e. over the kind of object expressions refer to within intensional contexts. Taking x as a variable ranging over 'normal' objects, α as a variable ranging over whatever referring expressions refer to in intensional contexts and Quine-quotes, '$\ulcorner$' and '$\urcorner$', to indicate the shift of reference modal operators achieve, one can come to recognize that:

(2.20) $\exists x$ necessary $\ulcorner x > 5 \urcorner$

still does not make much sense. (2.20), however, cannot immediately be derived from (2.12) by existential generalization, whereas (2.21) could:

(2.21) $\exists \alpha$ necessary $\ulcorner \alpha > 5 \urcorner$.

It be noted in passing that this approach accounts neatly for the semantic import of intermediate occurrences of expressions, a definite semantic relation being alleged to obtain between expressions when they have vulgar and intermediate occurrence.

Proceeding along these lines one could even try to dispense with quantifying over objects. Coreferentiality of two expressions can then be taken as some special relation between intensional objects. If so, it might seem that all opacity is then banished from the semantical framework. Quine, however, has demonstrated that one should watch one's step here. Intensional objects have an air of elusiveness and tend to remain intangible if left unanalysed. Without proper restrictions on their identity criteria the opacity troubles are immediately reinstated:

> For, where A is any intensional object, say an attribute, and 'p' stands for an arbitrary true sentence, clearly:
>
> (2.22) $A \equiv \imath x \cdot (p \wedge x \equiv A).$
>
> Yet, if the true sentence represented by 'p' is not analytic, then neither is (2.22), and its sides are no more interchangeable in modal contexts than are [...] '9' and 'the number of the planets'.
> (Quine [1961], p.152-3. Formal notation mine.)

One way of regaining substitutional transparency is by specifying the identity criteria for intensional objects such that any two ways of uniquely specifying that object are logically equivalent, i.e. for each intensional object x:

$$(\forall y(\varphi(y) \leftrightarrow y \equiv x) \wedge \forall y(\psi(y) \leftrightarrow y \equiv x)) \rightarrow \Box\forall y(\varphi(y) \leftrightarrow \psi(y)).$$

The price one has to pay for this move is, however, that your opponents can now argue along lines similar to those of the argument (1.9)–(1.14) on page 6 of the introduction, proving that modal distinctions collapse.

A third stance has been taken by Smullyan (Smullyan [1948]). His solution was in broad outline the one that has already been alluded to in the introduction: one should distinguish carefully between *de re* and *de dicto* readings of modal statements and the failure of substitution *salva veritate* can no longer be held against the principle of indiscernibility of identicals. Smullyan has pointed out Russell's theory of the contextual definition of descriptions can be applied to differentiate between *de re* and *de dicto* readings of modal statements. Modalities are here conceived of sentential operators be attached to open sentences and they will accordingly be depicted by $\Box$. (2.12) and 'The number of planets is nine' are then represented as respectively:

(2.23) $\Box 9 > 5$

(2.24) $\exists x(\forall y(y \text{ numbers the planets} \leftrightarrow x = y) \text{ and } x = 9).$

Just as within Russell's theory of descriptions

(2.25) The King of France is not bald,

could be analysed as either of the non-equivalent pair (2.26) and (2.27),:

(2.26) $\exists x$ (x is the King of France and $\neg$(x is bald))

(2.27) $\neg\exists x$ (x is the King of France and x is bald)

(2.15) is ambiguous between (2.28) and (2.29):

(2.28) $\exists x$ x is the number of planets and $\Box$ (x is greater than five), (2.29) $\Box\exists x$ (x is the number of planets and x is greater than five).

Moreover, he demonstrated how the true *de re* reading (2.28) can be derived from (2.23) and (2.24) within the system of *Principia Mathematica*, whereas as to (2.29) one's hopes are that this is not the case.

In Smullyan's analysis the failure of substitution *salva veritate* and existential generalization in the case of *de dicto* readings of modal statements, is accounted for by a well-known principle of classical first-order predicate logic: if a variable is bound by a quantifier it cannot be substituted, nor can it be bound by another more outside quantifier. On this account quantifiers play an indispensable role in the formalization of the *de re-de dicto* distinction. The important thing Smullyan has shown is how the *de re-de dicto* distinction and the feasibility of substitution can be conceived of as a result of variables being bound from within or without the scope of a modal operator.

Smullyan's arguments turn on a fundamental distinction between variables and proper names on the one hand, and expressions that can be defined contextually as definite descriptions, on the other. This distinction can be seen to foreshadow Føllesdal's between singular and general terms, be it that on Smullyan's account general terms are analysed away. Still, with $\Box$ allowing for variables being bound within its scope from without, Smullyan's solution will not succeed in removing Quine's doubts concerning the metaphysical purity of quantified modal logic. In (2.28) there still is a variable in the scope of a modal operator that is to be bound by an outside quantifier. It was precisely the intelligibility of constructions like this Quine took so much pain to dispute.

It should be noted, however, that Quine's arguments that were meant to refute quantified modal logic were directed merely at *de dicto* modality. In a reply to Kaplan, Quine was frank to admit this:

> In my treatment of belief I distinguished between an opaque and a transparent version, but in modal logic I got no further that the opaque. I agree with Kaplan was thus asymmetric and that the fact of the matter is symmetric. The distinction between opaque and transparent on the modal side is the distinction between what Chisholm, reviving scholastic terminology, calls *necessitas de dicto* and *necessitas de re*. [...] [M]y treatment of modal logic was brief and negative: I was content to outline the opacity troubles.
> (Davidson and Hintikka [1969], p.343-4)

This passage suggests that Quine took opacity, on the one hand, and modalility *de re* and *de dicto*, on the other, as separate issues.

Taking cue from his treatment of propositional attitudes in Quine [1956], Quine restored the symmetry between his treatment of propositional attitudes and modal contexts in a later paper on intensional contexts (Quine [1977]). The basic idea is there that one can disentangle referential use and non-referential use of expressions occurring in modal contexts in much the same manner as use and mention of expression can be separated in (2.4) and (2.5) by rewriting these as respectively (2.10) and (2.11). Just as (2.4) is in an important sense about Giorgione, (2.15), if understood in its *de re* reading, is about the actual number of planets, i. e. nine. Still, modal contexts are liable to give rise to opacity phenomena. Both these features of *de re* modalities could be made explicit in a formal notation. Being attentive to not getting immersed too deeply in modal logic, Quine persists in anatomizing modalities as a (semantical) predicate, 'Nec'. 'Nec', however, differs from the semantical predicate by the same name in Quine [1953c] in the sense that it now gets *multigrade status* (not my fault), i. e. for any $n > 0$ 'Nec' is an $n+1$-ary predicate holding between the name of an n-ary predicate and n objects. Treating propositions as zero-place predicates, his original 1953 'Nec' is then the borderline case in which $n = 0$.[6] *De re* and *de dicto* readings can now be accommodated as follows:

> When predication in the mode of necessity is directed upon a variable, the necessity is *de re*: the predicate is meant to be true of the value of the variable by whatever name, there being indeed no name at hand. 'Nec('odd',x)' says of the unspecified object x that oddity is of its essence. Thus it is true not only that Nec('odd',9), but equally that Nec('odd',number of planets), since this very object 9, essence and all, happens to *be* the number of planets. The 'Nec' notations accommodates *de dicto* necessity too, but differently: the term concerned *de dicto* is within the quoted sentence or predicate. Thus 'Nec '9 is odd'', unlike 'Nec('odd',9)', is *de dicto*, and 'Nec('number of planets is odd')', unlike 'Nec('odd', number of planets)', is false. (Quine [1977], p.269)

In doing so Quine has detached the opacity from the modal predicate and has shifted it to the predicate in the first argument place. Note that the predicate in the first argument place occurs within quotation marks 'Nec' itself is entirely transparent. Any referential expression occupying one of the argument places can be substituted for a coreferential one *salva veritate*. This shows how one can distinguish between *de re* and *de dicto* and generalize existentially over referential expressions if they get a *de re* reading without committing the error of quantifying

[6] Quine's treatment does not run entirely parallel to the Giorgione case. This cannot be used against him since the comparison in this form is not his but mine. Yet, by quoting predicates, rather than whole statements Quine forestalled the possibility of referring expressions occurring in both a referentially transparent, and in a substitutionally opaque position (cf. introduction). If so, the comparison of *de re* modality with (2.4) in which the use and mention of 'Giorgione' were confused goes awry. In 'Nec('odd', number of planets)', 'number of planets' does not recur anywhere in an opaque context. Quoting whole sentences 'The number of planets is odd', if taken *de re*, would become 'Nec('The number of planets is odd', number of planets)'. A reason to opt for quoting predicates rather then whole sentences, is that it might be rendered obscure what is said about which object if a referring expression in an argument place of 'Nec' is replaced by a coreferential one. What is said about nine in, e. g. 'Nec('The number of planets is equals $\sqrt[3]{729}$', nine)'? Does it say of the number of planets, i e. nine, that it is necessarily equals $\sqrt[3]{729}$, which is true, or about $\sqrt[3]{729}$ that whatever the number of planets may be the latter necessarily equals the former, which would be false? By using predicates one can avoid any such confusion. Quine proposes the use of a predicate forming device $\ni$. Then '$xy\ni\varphi(x,y)$' is a two place predicate, with the additional feature that its argument places are then made explicit. The ambiguity in the former case can then be resolved by having 'Nec('$x\ni x$ equals $\sqrt[3]{729}$', nine)' represent the first reading and 'Nec('$x\ni$the number of planets equals x', nine)' the second. Yet, it cannot be difficult to achieve a similar result for names of sentences. Dagger-quotes ($^\dagger\ldots{}^\dagger$) could be defined as a name-forming device applicable to sentences and also allows for occupied argument places to be labelled. Let for instance †1:The number of planets. equals $\sqrt[3]{729}^\dagger$be a name of the sentence 'The number of planets equals $\sqrt[3]{729}$', with the position of 'the number of planets' being labelled. Then ambiguity of the example could then be eliminated by distinguishing between 'Nec(†1:The number of planets. equals $\sqrt[3]{729}^\dagger$, nine)' and 'Nec(†The number of planets equals 1:$\sqrt[3]{729}^\dagger$, nine)'.

into an opaque context. This is not so much as saying that Quine's anxiety with respect to modal logic are to be brushed aside. He still views 'Nec' with a suspicious eye:

> The reconstruction of '$\Box$' in terms of 'Nec' has lent some clarity to the foundations of modal logic by embedding it in extensional logic, quotation, and a special predicate. Incidentally the contrast between *de re* and *de dicto* has thereby been heightened. But the special predicate takes some swallowing. In its monadic use it is at best the controversial semantic predicate of analyticity, in its polyadic use it imposes an essentialist metaphysics. Let me be read, then, as expounding rather than propounding. I am in the position of a Jewish chef preparing ham for a gentile clientele. Analyticity, essence, and modality are not my meat. (Quine [1977], p.269-70)

The Fregean way of tackling the problem of substitution in intensional contexts bears a striking resemblance to Quine's point that one cannot substitute within quotation marks in one important respect. Just as a term between quotation marks refers to a name of the object it normally denotes, a referential term denotes its sense in an intensional context on a Fregean account. In both cases replacing the term by another that in extensional contexts denotes the same object, would not in general be a case of replacing terms that are coreferential when they occur within quotation marks or in an intensional context. One object may have different names and coreferential terms may have different senses. Accordingly, the alleged counterexamples against Leibniz's Law above fail as such for the Fregean since the terms that are substituted for one another occur within an intensional context where they do *not* denote the same object. Quine would point out that substituting coreferential terms in intensional contexts will eventually be tantamount to substituting expressions within quotation marks. This will come to be recognized when use and mention are properly disentangled in an equivalent extensional statement.

This emphasis on the reference of terms when they occur in intensional contexts, however, seems slightly beside the point. The case is analogous to a similar problem involving such determiners as 'all'. If the Aristotelian dogma of the subject-predicate structure of sentences is adhered to, the reference of 'each man' in:

(2.30) Each man is mortal

might at first sight seem to be essential to the truth-conditions of (2.30). Questions could be raised as to the denotation of 'each man'; whether it denote the class of all men, an arbitrary man, the concept of man or, perhaps, the universal man, if any. By dismissing the subject-predicate dogma and analysing (2.30) as:

(2.31) $\forall x(x \text{ is a man} \rightarrow x \text{ is mortal})$,

modern quantified predicate logic does efficiently away with such conundrums concerning the denotation of 'each man'. The point I wish to make here is that in a similar manner, the behaviour of referential terms in the scope of intensional operators should be understood when they get a *de dicto* reading. What is relevant to the truth conditions of, e. g., in (2.15) when understood in a *de dicto* fashion, not so much the reference of 'the number of planets' as it occurs in (2.15), but rather the properties the objects denoted by 'the number of planets' in the various possible worlds have in the respective possible worlds. The standard semantical interpretation of the universal quantifier is such that, in order to determine the truth value of $\forall x\varphi$, one has to evaluate φ with respect to each value for x. In a similar fashion, in order to assess a modal statement, one has to determine the truth-value of the formula in the scope of the modal operator with respect to all possible worlds and the denotations of the *de dicto* terms in those worlds. As such it is the denotation of the *de dicto* terms in other possible worlds and the properties they have over there that is relevant to the truth-value of modal formulae, rather than the denotation of the referential expressions when they occur in the scope of a modal operator. It is also in this sense that a modal operator can be said to *bind* terms in its scope.[7]

Since it is the various objects a referential term t that is bound by a modal operator may denote in other possible worlds that are relevant to the truth-value of the sentence in which t occurs, one cannot expect a substitution of a term t with another term that merely denote the same object in the actual world not to precipitate a change of truth value. Bound terms cannot be substituted and the failure of substitutivity *salva veritate* of coreferential terms in modal contexts can likewise be explained.

On this conception, *de re* readings of referential terms in modal contexts can be accounted for by their not being bound by the modal operator in question. As such, these terms can be substituted by coreferential ones without need for further explanation. The reference of a term that is not bound by a modal operator, is the value it obtains from the world in which the respective modal sentence is evaluated. This account of *de re* readings of terms bears some important similarities with Kaplan's account of directly referential terms, about which more in chapter 5, below.

Rather than being particularly original, many of the above remarks are tacitly assumed in the practice of quantified modal logic. The idea of analysing the difference between *de re* and *de dicto* readings of modal statements as a binding phenomenon had already been present in Smullyan's exposition. On the latter's proposal, however, the binding was achieved by the quantifiers and his analysis was accordingly restricted to the variables. Russell's theory of descriptions, however, made that other referential terms could also be dealt with. Notwithstanding the power and popular appeal of the theory, some mysteries remain. As Føllesdal has pointed out the modal operators should be transparent for the variables in order to account for quantifying in but opaque for general ones to prevent modal distinctions from collapsing. The quantifiers only bind variables, however. Consequently, Smullyan's account leaves the failure of substitutivity of general terms, when not analysed away in the Russellian manner, unaccounted for.

The proposal that will be elaborated on in the sequel, is to retain the spirit of Smullyan's analysis by conceiving of the difference between *de dicto* and *de re* as a binding phenomenon but to have the modal operators themselves take over the task allotted to the quantifiers by Smullyan: the binding of terms. Modal operators are

[7] This is not so much as saying that a Fregean sense is banished from the semantical framework altogether. Bound terms do not loose their descriptive content, if any. A Fregean sense may be very relevant to determining the references of the bound terms in the various relevant possible worlds. Accordingly, the contention that modal operators bind terms in their scope does not run counter to a Fregean semantics as such.

thought of as being able to bind any kind of term, be it a variable, an individual constant or a complex functional term. The formal modal languages with which the remainder of this thesis is concerned can differ with respect to the terms they allow the modal operators to bind. With each language a set of indices consisting of sets of terms will be specified with which the modal operator can be indexed. Just as the universal quantifier $\forall$ can be specified to bind different variables (both $\forall x\varphi$ and $\forall y\varphi$ are possible), $\Box$ can get indexed with any of the sets of terms in the set of indices. So if Z and Y are indices of some modal language which also has φ as a formula then both $\Box_Z\varphi$ and $\Box_Y\varphi$ are syntactically well-formed formulae. In the symbolism adopted throughout this thesis, modal operators are indexed with the terms they do *not* bind, i. e. for which they are (substitutionally) transparent and for which they enforce a *de re* reading. The ambiguity of (2.15), as noted by Smullyan, will thus be made explicit in the following manner:

(2.32) $\Box_{\{\text{the number of planets}\}}$ the number of planets is greater than five

(2.33) $\Box_{\emptyset}$ the number of planets is greater than five.

The sensitivity of modal operators to predicate logical structure of the formulae in their scope will thus be accounted for. Importantly, the set of all terms is not in general ineligible as an index. Still, in defiance of Quine's prophecies, no collapse of modal distinctions is immanent if such languages, which contain indexed modal operators, are provided a suitable semantics. The next chapter is devoted to this.

Chapter 3

Semantics for Indexed Modal Operators

In this chapter a Kripke-style semantics for a class of modal predicate logical languages $\mathcal{L}$ will be presented along with some formal results. The models used in the semantics are very much like the common Kripke models for quantified modal logic. Yet, the languages of $\mathcal{L}$ differ from the usual quantified modal languages in that they do not contain the quantifiers $\forall$ and $\exists$ and in that the modal operators are indexed by a set of terms. This may raise an occasional eyebrow, for was it not exactly the interaction between the quantifiers and the modal operators that granted quantified modal logic its initial interest? In chapter 1 it was emphasized that the predicate logical structure of formulae is relevant to the formal analysis of modalities, even if the quantifiers are absent. The distinction between *de re* and *de dicto* readings of modal statements was taken as an indication of this. In the previous chapter I argued that the distinction between *de re* and *de dicto* readings of referential terms in modal contexts could be accounted for by conceiving of modal operators as devices that can bind terms in their scope. The appreciation of the *de re*-*de dicto* distinction as basically a binding phenomenon had already been present in Smullyan [1948]. Smullyan demonstrated how *de re* and *de dicto* could be analysed in terms of the relative scope of quantifiers and modal operators. Still, his examination did not extend to referential expressions that are not analysed as definite descriptions.

In the modal predicate logic that will be advanced here, **PLuM**, the modal operators themselves are bestowed a sensitivity to the predicate logical structure of the formulae in their scope. The underlying idea is that the principles of modal logic are not in essence of a propositional nature only. Moreover, the difference between *de re* and *de dicto* readings of modal statements are not thought of as being merely spin-off to the interaction between modal operators and quantifiers. The *de re*-*de dicto* distinction is rather thought of as indicative of there being essentially predicate logical principles of modal logic. The modal predicate logic **PLuM** is presented in an effort to make these principles explicit.

The languages of $\mathcal{L}$ comprise functional terms, which, as it turns out, tend to complicate semantical matters considerably. There are two main reasons to include them in the analysis nonetheless. First, the languages of $\mathcal{L}$ do not contain explicit quantifiers and as such lack the means to deal with definite descriptions in a Russellian manner. The behaviour of definite descriptions in modal contexts, however, is particularly relevant to issues relating to the *de re-de dicto* distinction. Had the languages of $\mathcal{L}$ been short of any means to account for them, **PLuM** would also have been a poor tool for anyone interested in the formal semantics of natural language. The second consideration is more congenial to the quintessence of this paper. In **PLuM** Quine's modal collapse can be averted without stipulating the modal operators being transparent for individual constants and variables and opaque for others, such as functional terms. Without functional terms, or similar devices, Quine's paradox could not even be formulated and the point would be lost. The semantics of **PLuM** are such that the modal operators need not bind the functional terms in their scope neither should the constants and the variables always remain free. Still, modal distinctions can be accounted for.

Smullyan dealt with the difference between *de re* and *de dicto* readings by having the quantifiers bind variables from within and from without the scope of the modal operators. In **PLuM** the modal operators themselves attend to the binding of terms in their scope and in this manner they account for the *de re-de dicto* distinction. It might be wondered how similar these approaches are and to what extent the notions of a quantifier binding a variable and a modal operator binding a term are alike. Chapter 6, below, which will concern the modal foundations of quantification, may provide an answer. As matters evolve, quantifiers may be conceived of as a special kind of indexed modal operator.

3.1 Elements of the Modal Predicate Logic PLuM

3.1.1 The Languages of $\mathcal{L}$

The class of modal languages in which the investigations of this thesis will be conducted, $\mathcal{L}$, differ only slightly from usual formal languages for modal predicate logics. Each language L of $\mathcal{L}$ contains modal operators each of which is indexed with a, possibly empty, set of terms. The quantifiers $\forall$ and $\exists$, however, do not rank among the primitive symbols. The lexicon L of each language $L \in \mathcal{L}$ is specified as follows:

Definition 3.1 (The lexicon of a language $L \in \mathcal{L}$)
The lexicon L of a language L of $\mathcal{L}$ consists of:

(*i*) A non-logical part. For each $n \in \omega$:

- a countable set, $\textsc{Func}_L$, of n-ary function symbols $f_1^n, f_2^n, \ldots$
- a countable set, $\textsc{Rel}_L$, of n-ary relation symbols $R_1^n, R_2^n, \ldots$

Since n is allowed to be zero, the set of individual constants, $\textsc{Cons}_L$, is defined as the set of zero-place function symbols. Zero-place relation symbols make up the set of propositional variables.

(*ii*) A countable set, $\textsc{Var}_L$, of variables $v_1, v_2, \ldots$

(*iii*) the logical logical constants: $\perp$ (falsum), $\rightarrow$ (material implication), $\equiv$ (identity), and the modal operator $\Box$ ('necessary') (*iv*) parentheses '(' and ')'.

In the sequel a, b, c; x, y, z, v, w and f, g, h, with or without natural numbers as indices or with $'$, $''$ as superscripts, are used as meta-variables ranging over respectively individual constants, variables and function symbols. As meta-variables ranging over propositions, unary, and n-ary relation symbols respectively p, q, r; P, Q and R, S will usually be employed.

For each language $L \in \mathcal{L}$ the set of terms, TERM_L is defined inductively as usual.

Definition 3.2 (Terms of L, TERM_L) For each $L \in \mathcal{L}$ the set of terms, TERM_L, is the smallest set such that:

(*i*) $\text{CONS}_L \subseteq \text{TERM}_L$

(*ii*) $\text{VAR}_L \subseteq \text{TERM}_L$

(*iii*) If f is an n-ary function symbol in FUNC_L and $t_1, \ldots, t_n \in \text{TERM}_L$ then $f(t_1, \ldots, t_n) \in \text{TERM}_L$

$t_1, t_2, \ldots$ and $t, t', t'', \ldots$ will be used as the meta-variables ranging over terms. The distinction between variables and constants is for the greater part customary. The logic for $\mathcal{L}$ that will be presented in the next section makes variables and individual constants of a language $L \in \mathcal{L}$ behave in much the same manner, it will be convenient to refer to them as a single category of atomic terms, ATTERM_L. Thus, $\text{ATTERM}_L = \text{VAR}_L \cup \text{CONS}_L$.

For each language L of $\mathcal{L}$ a non-empty set of indices, INDEX_L, is defined as a set of subsets of TERM, i.e. $\text{INDEX}_L \subseteq \wp(\text{TERM})$.

Apart from the clause for $\Box$, the syntax of the formulae of any language $L \in \mathcal{L}$, FORM_L, is much as one would expect:

Definition 3.3 (Formulae of L, FORM_L.)
The set of formulae of any language L in the class $\mathcal{L}$ is the smallest set such that;

(*i*) If R is an n-ary relation symbol in REL_L and $t_1, \ldots, t_n \in \text{TERM}_L$ then $R(t_1, \ldots, t_n) \in \text{FORM}_L$.

(*ii*) If $t_1, t_2 \in \text{TERM}_L$ then $t_1 \equiv t_2 \in \text{FORM}_L$.

(*iii*) $\bot \in \text{FORM}_L$

(*iv*) If $\varphi, \psi \in \text{FORM}_L$ then $(\varphi \to \psi) \in \text{FORM}_L$

(*v*) If $\varphi \in \text{FORM}_L$ and $Z \in \text{INDEX}_L$ then $(\Box_Z \varphi) \in \text{FORM}_L$

N.B.: Also here subscripts will be omitted if it is unequivocal which language is concerned. Parentheses will likewise be omitted in cases in which no ambiguity can arise.

The Greek lower-case letters φ, ψ, χ, with or without natural numbers as indices, will usually be employed as meta-variables ranging over formulae of any language L of $\mathcal{L}$. $\top$, $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \leftrightarrow \psi$, $\Diamond_X \varphi$ abbreviate respectively $\bot \to \bot$, $\varphi \to \bot$, $\neg(\varphi \to \neg\psi)$, $\neg(\neg\varphi \wedge \neg\psi)$, $(\varphi \to \psi) \wedge (\psi \to \varphi)$, and $\neg\Box_X \neg\varphi$, as is common practice.

The languages of $\mathcal{L}$ can be distinguished by their lexicons or by their respective sets of indices. Most of this study will be restricted to the class of languages

$\mathcal{L}^* \subseteq \mathcal{L}$, for which the indices are closed under taking subterms. Certain results can only be obtained for a subclass of $\mathcal{L}^*$. Each of the indices of the languages in this class, $\mathcal{L}^{**}$, is either a subset of the atomic terms or the set of all terms. Let further $\mathbf{L}^*$ and $\mathbf{L}^{**}$ the largest languages of respectively $\mathcal{L}^*$ and $\mathcal{L}^{**}$.

Definition 3.4 Define $\mathcal{L}^*$, $\mathcal{L}^{**}$, and for each lexicon L $\mathbf{L}^*$ and $\mathbf{L}^{**}$ as:

(i) $\mathbf{L}^*$ is that language in $\mathcal{L}$ with L as lexicon such that $\text{INDEX}_{\mathbf{L}^*} = \{Z \subseteq \text{TERM} \,|\, \text{Z is closed under subterms}\}$

(ii) $\mathbf{L}^{**}$ is that language in $\mathcal{L}$ with L as lexicon such that $\text{INDEX}_{\mathbf{L}^{**}} = \wp(\text{ATTERM}) \cup \{\text{TERM}\}$

(iii) $\mathcal{L}^* = \{L \in \mathcal{L} \,|\, \text{for some lexicon L: } L \subseteq \mathbf{L}^*\}$

(iv) $\mathcal{L}^{**} = \{L \in \mathcal{L} \,|\, \text{for some lexicon L: } L \subseteq \mathbf{L}^{**}\}$

Clearly, $\mathcal{L}^{**} \subseteq \mathcal{L}^*$. The restriction to languages which indices are closed under taking subterms is a natural one both from a philosophical and a technical aspect. This will become clear when the semantics for **PLuM** has been presented.

Another important syntactic concept is that of a referential (occurrence of) a term in a formula. A term t occurs referentially in a formula φ if it is not in the scope of a modal operator $\Box_Z$ with $t \notin Z$, i. e. if t is bound by no such $\Box_Z$ in φ. Eventually, it will be these referential occurrences of terms that obtain a *de re* interpretation in the semantics here presented. This concept should not be taken as were it to capture Quine's notion of a 'purely referential' occurrence of a term. The logic for the languages of $\mathcal{L}$ preserves the distinction between a term occurring referentially and a term being substitutable for any coreferential one, which sometimes threatens to get confused in Quine's informal reflections. The interactions between referentiality and substitutivity are more intricate and will be dealt in section 3.2.3 of this chapter.

Definition 3.5 (Referring Terms)

- $\text{RT}(x) := \{x\}$
- $\text{RT}(c) := \{c\}$
- $\text{RT}(f^n(t_1, \ldots, t_n)) := \bigcup_{1 \le i \le n} \text{RT}(t_i) \cup \{f^n(t_1, \ldots, t_n)\}$
- $\text{RT}(R^n(t_1, \ldots, t_n)) := \bigcup_{1 \le i \le n} \text{RT}(t_i)$
- $\text{RT}(t_1 \equiv t_2) := \text{RT}(t_1) \cup \text{RT}(t_2)$
- $\text{RT}(\bot) := \emptyset$
- $\text{RT}(\varphi \to \psi) := \text{RT}(\varphi) \cup \text{RT}(\psi)$
- $\text{RT}(\Box_X \varphi) := \text{RT}(\varphi) \cap X$

3.1.2 Semantics for the Languages of $\mathcal{L}$

The semantics for **PLuM** (**P**redicate **L**ogic **u**n-quantified and **M**odal) will be given by interpreting the languages of $\mathcal{L}$ on possible words models, which are very much like Kripke models, except that the assignment and interpretation functions are integral parts of the possible worlds themselves rather than of the model as a whole.[1] The possible worlds are taken to be full predicate logical models. As a consequence, the variables behave in much the same way as the individual constants.

Definition 3.6 ($\textsc{Ass}_D$) Let for each set of objects D, $\textsc{Ass}_D$ be the set of (total) functions $g : \textsc{Var} \to D$.

Definition 3.7 (Possible worlds and m-models for the languages $L \in \mathcal{L}$)

(i) a *(possible) world* w is a triple $\langle D_w, [\![\cdot]\!]_w, g\rangle$ with:

- D_w a set of objects
- $[\![\cdot]\!]_w$ an interpretation function for the non-logical constants such that:
  - for each individual constant $c \in \textsc{Cons}_L$: $[\![c]\!]_w \in D_w$
  - for each n-ary function symbol $f \in \textsc{Func}_L$:
    $[\![f]\!]_w$ is a total function $\mathbf{h} : (\bigcup_{w' \in \mathcal{W}} D_{w'})^n \to \bigcup_{w' \in \mathcal{W}} D_{w'}$
    such that if $d_1, \ldots, d_n \in D_w$ then $\mathbf{h}(d_1, \ldots, d_n) \in D_w$[2]
  - for each n-ary relation symbol R : $[\![R]\!]_w \subseteq D_w^n$
- $g \in \textsc{Ass}_{D_w}$
  N.B.: In the sequel of this thesis the notation $\textsc{Ass}_w$ for Ass_{D_w} will be adhered to.

(ii) Define an m-model $\mathcal{M}$ as a pair $\langle \mathcal{W}, \mathcal{R}\rangle$, with:

- $\mathcal{W}$ is a set of (possible) worlds
- $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$

The objective of **PLuM** is to define the notions of *de re* and *de dicto* modality in a mathematically precise way and thus rendering them feasible for the formal sciences[3] as well as in the formal analysis of natural language. The notion of *de re* and *de dicto* modality that the indexed modal operators are to capture is restricted to the interpretation of the terms. The semantics of **PLuM** is constructed along

[1] As for the interpretation function for the non-logical lexicon this is merely a matter of definition. The interpretation function could also be taken as a function from possible worlds to functions with the non-logical lexicon as its domain and its various interpretations as its range. The same procedure could be followed with respect to the assignment function. A definition along these lines would render the notion of a frame, see below, more transparent. But even so, one would be departing from the common practice to assign variables their values without reference to the possible worlds. Here variables are assigned their values for each possible world separately, and consequently they behave as were they individual constants. The merely syntactic difference between variables and individual constants is maintained, in order to be able to accommodate referential appliances such as λ's and quantifiers more elegantly in the later stages of this thesis.

[2] As an alternative definition the following might seem feasible:
– for each n-ary function symbol f : $[\![f]\!]_w$ is a total function $\mathbf{h}$: $D_w^n \to D_w$.
In the development of **PLuM**, however, it will become apparent that in the course of evaluating a modal formulae, terms may in some cases get interpreted in a world (or, to be more precise, a wur) outside the domain of that world. If any such term, say t, were to occur in an argument position of a functional symbol f, it may happen that the functional term $f(t_1, \ldots, t, \ldots, t_n)$ will become undefined. This would force one into a position where one has to make choices how to treat undefined terms.

[3] At this point the question which applications the notions of *de re* and *de dicto* could have in the formal sciences becomes urgent. I am afraid the philosophical introduction of this master thesis is deficient in this respect.

such lines that a modal formula $\Box_Z\varphi$ is interpreted in such a way that precisely those terms t such that $t \in RT(\varphi)$ get a *de re* reading.

The salient feature of *de re* modality is that it concerns the properties *objects* denoted by referring expressions in the actual world, have in other possible worlds. On the other hand, if a term t obtains a *de dicto* interpretation in a modal sentence φ, the reference of t in each of the accessible possible worlds separately is relevant to the truth-value of φ. The denotations of t in the various worlds may widely differ. In short, if a term t is to get a *de re* interpretation, only one object is relevant, *viz.* the object t refers to in the actual world. Whereas, if t were to get a *de dicto* reading the objects t refers to in the various accessible possible worlds are of concern to the the truth-value of the sentence in which t occurs. Thus the basic idea underlying the semantics of **PLuM** is that a term t occurring in the scope of a modal operator $\Box_Z$ with $t \in Z$ should get the interpretation it gets in the world in which the modal sentence is evaluated rather than its interpretations in the various worlds accessible from that world. Terms that have a *de re* occurrence should, roughly speaking, derive their interpretation from the world the sentence containing all the modal operators in which scope it lies, is evaluated. If, however, $t \notin Z$, t will have to be interpreted in the worlds accessible from the one in which the evaluation takes place. An example may help to clarify matters:

Example 3.8 Consider a model $\mathcal{M}=\langle\mathcal{W},\mathcal{R}\rangle$ with $\mathcal{W}$ containing two worlds w_1 and w_2 and the latter being accessible from the former only, i. e. $w_1\mathcal{R}w_2$. Assume $[\![t]\!]_{w_1} = d$, for some term t.Suppose further that one should evaluate the formula $\Box_{\{t\}}Pt$ in w_1. t should then get a *de re* interpretation. As such, one is interested in whether $d \in [\![P]\!]_{w_2}$, i.e. whether $[\![t]\!]_{w_1} \in [\![P]\!]_{w_2}$. Whether $[\![t]\!]_{w_2} \in [\![P]\!]_{w_2}$ would be of interest only if one were to evaluate the *de dicto* statement $\Box_{\emptyset}Pt$ in w_1. So, whereas to the evaluation of $\Box_{\{t\}}Pt$ the interpretation of t in w_1 is relevant, it is, among others, the denotation of t in w_2 that matters in the case of $\Box_{\emptyset}Pt$.

In most better-known Kripke-style semantics for modalities the semantic interpretation of $\Box$ is usually defined along the following lines:

$$\mathcal{M}, w \models \Box\varphi \Longleftrightarrow \text{ for all } w' \text{ such that } w\mathcal{R}w' : \mathcal{M}, w' \models \varphi.$$

In the course of evaluating a modal sentence in a world w one may be forced to investigate other possible worlds. Yet, the definition above, at least as it stands, has as a consequence that it is utterly immaterial to the assessment of the truth value of a formula in a world at which stage of the evaluation process of a formula that assessment takes place. Consider again example 3.8. In establishing the truth of the *de re* modal formula $\Box_{\{t\}}Pt$ in w_1, one had to assure oneself of $[\![t]\!]_{w_1}$ being in the extension of P in w_2, i. e. whether $[\![t]\!]_{w_1} \in [\![P]\!]_{w_2}$. The question is how this can be achieved if, in the course of the evaluation process, one loses the information the truth-value of which formula in which world one initially started out to investigate. One could proceed by imposing the restriction on the models that the interpretation of a term in a possible world should coincide with the interpretation of that term in any other of the accessible possible worlds. The catch of this approach, however, is that, if this restriction is to apply to all terms of the language and the language contains sufficiently strong term-forming devices such as $\imath$-operators or function symbols, a semantical variant of Quine's argument can be formulated. If one should ensure that for each n-place relation symbol R, and all terms $a, t_1, \ldots, t_n$ that the interpretations of both $\imath x \cdot x \equiv a \wedge t_1 \equiv t_2$ and $\imath x \cdot x \equiv a \wedge R(t_1, \ldots, t_n)$ are the same in each possible world, it can easily be shown that one ends up with models in which any two possible worlds of which one is accessible from the other will coincide on the interpretation of the terms and relation symbols.

In order to avoid this outcome, one could divide the class of terms into rigid and non-rigid designators. The rigid designators are then usually taken to be the variables and the individual constants. This, of course, corresponds to Føllesdal's distinction between singular and general terms alluded to in the introduction and several other places above.

The line taken in **PLuM** is quite different. The fundamental idea underlying it is to refrain from denying the feasibility of certain of the models as defined in definition 3.7. The restrictions on the interpretation of terms are rather imposed during the process of evaluating a modal sentence. The restrictions on the interpretation of the terms in the possible worlds are made dependent on the modal formulae and the worlds in which they are evaluated. This is achieved by means of quite a simple semantical mechanism that is controlled by the indices of the modal operators occurring in the formulae in question. When, for instance, assessing the truth-value of $\Box_{\{t\}}Pt$ in w_1 of example 3.8 above, w_2 will be examined with respect to the truth of Pt with the proviso that t get the interpretation it has in w_1. This restriction on the interpretation of t in w_2 is imposed when determining the truth value of $\Box_{\{t\}}Pt$ in w_1, because t occurs in the index of the modal operator of $\Box_{\{t\}}Pt$. When assessing $\Box_{\emptyset}Pt$ in w_1 the interpretation of the terms in w_2 will not be curbed thus. Formulae and terms are thus no longer interpreted in the worlds of the model but rather in the world under a certain restriction on the interpretation of the terms (wurs). This device imposing restrictions on the interpretation of terms is not very much unlike the way assignments are manipulated when evaluating quantified formulae in standard quantified predicate logic. It also turns out that there is no longer a need for creating dichotomies between singular and general terms in order to avert the collapse of the number of distinguishable possible worlds.

The idea is simple enough. The restrictions on the interpretation of terms in a world w' are determined by the indices of the (outermost) modal operators of the formula evaluated in a world w from which w' is accessible. The corresponding definition of a wur, however, is slightly complicated by the possibility of modal operators being nested, which makes that inductive definitions are in order. An intimate connection exists between the indices of the modal operators and the restrictions on the interpretation of terms. Since the set of indices is particular to a language $L \in \mathcal{L}$, only a limited number of restrictions are relevant to the evaluation of formulae of a specific language in the m-models. Hence it is convenient to define restrictions relative to a language L of $\mathcal{L}$.

Definition 3.9 (Restrictions and Worlds under Restriction) Define inductively for each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and each L in $\mathcal{L}$, the set of *restrictions on the interpretation of terms* ($\textsc{rest}^L_{\mathcal{M}}$) and the set of *worlds under a restriction*, wurs, ($\textsc{wur}^L_{\mathcal{M}}$) simultaneously as the smallest sets such that:

(i) [-] $\in \textsc{rest}^L_{\mathcal{M}}$

(*ii*) $\{w[\text{-}] \mid w \in \mathcal{W}\} \subseteq \text{WUR}^L_{\mathcal{M}}$

(*iii*) if $X \in \text{INDEX}_L$ and $\varpi \in \text{WUR}^L_{\mathcal{M}}$ then $[X/\varpi] \in \text{REST}^L_{\mathcal{M}}$

(*iv*) if $w \in \mathcal{W}$ and $[\varrho] \in \text{REST}^L_{\mathcal{M}}$ then $w[\varrho] \in \text{WUR}^L_{\mathcal{M}}$

N.B.: The empty restriction [-] will usually be omitted as will the subscript $\mathcal{M}$ and the superscript L whenever no confusion is likely.

Set-theoretically a wur $\varpi \in \text{WUR}^L_{\mathcal{M}}$, where $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ could be construed as a tuple $\langle D_\varpi, [\![\cdot]\!]_\varpi \rangle$, with $D_\varpi \subseteq \bigcup_{w \in \mathcal{W}} D_w$ and $[\![\cdot]\!]_\varpi$ as an interpretation function that assigns to each term t an object $d \in D_\varpi$ and to each n-place relation symbol a set of n-tuples $\langle d_1, \ldots, d_n \rangle \in D^n_\varpi$. Accordingly the restrictions of definition 3.9 can be taken to be functions from worlds to wurs in postfix notation. The following definitions lay down the notions of D_ϖ and $[\![\cdot]\!]_\varpi$ in a rather more precise manner:

Definition 3.10 (The domain of a wur). Define for each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and each wur $w[\varrho] \in \text{WUR}$, $D_{w[\varrho]}$, inductively as:

- $D_{w[\text{-}]} := D_w$
- $D_{w[Z/\varpi']} := D_w \cup \{[\![z]\!]_{\varpi'} \mid z \in Z\}$

The interpretation of the terms of any language $L \in \mathcal{L}$ can now be defined relative to an m-model and a wur.:

Definition 3.11 (Interpretation of terms in a world under a restriction on the value of terms (wurs), $[\![t]\!]_\varpi$)
Let $w = \langle D, [\![\cdot]\!]_w, g \rangle$. Define for each language $L \in \mathcal{L}$, and each $t \in \text{TERM}_L$ and any m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $w[\varrho] \in \text{WUR}^L_{\mathcal{M}}$ inductively:

- $t \approx c \in \text{CONS}: \quad [\![c]\!]^{\mathcal{M}}_{w[\varrho]} := \begin{cases} [\![c]\!]^{\mathcal{M}}_\varpi & \text{if } [\varrho] = [Z/\varpi] \text{ and } c \in Z \\ [\![c]\!]_w & \text{otherwise} \end{cases}$
- $t \approx x \in \text{VAR}: \quad [\![x]\!]^{\mathcal{M}}_{w[\varrho]} := \begin{cases} [\![x]\!]^{\mathcal{M}}_\varpi & \text{if } [\varrho] = [Z/\varpi] \text{ and } x \in Z \\ g(x) & \text{otherwise} \end{cases}$
- $t \approx f(t_1, \ldots, t_n):$
  $$[\![f(t_1,\ldots,t_n)]\!]^{\mathcal{M}}_{w[\varrho]} := \begin{cases} [\![f(t_1,\ldots,t_n)]\!]^{\mathcal{M}}_\varpi & \text{if } [\varrho] = [Z/\varpi] \text{ and} \\ & f(t_1,\ldots,t_n) \in Z \\ [\![f]\!]_w([\![t_1]\!]^{\mathcal{M}}_{w[\varrho]}, \ldots, [\![t_n]\!]^{\mathcal{M}}_{w[\varrho]}) & \text{otherwise} \end{cases}$$

Employing this notion of a world under a restriction, the following definition monitors the satisfaction conditions of formulae $\varphi \in \text{FORM}_L$ $(\mathcal{M}, \varpi \models \varphi)$ in wurs $\varpi \in \text{WUR}^L_{\mathcal{M}}$ of m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$:

Definition 3.12 (Satisfaction in m-models) $\models$ determines the truth of formulae in possible worlds under a restriction. For each $L \in \mathcal{L}$, for each $\varphi \in \text{FORM}_L$ and for any m-model $\mathcal{M}=\langle \mathcal{W}, \mathcal{R} \rangle$, and any $w[\varrho] \in \text{WUR}^L_{\mathcal{M}}$ define inductively:

- $\mathcal{M}, w[\varrho] \models R(t_1, \ldots, t_n) :\iff \langle [\![t_1]\!]_{w[\varrho]}, \ldots, [\![t_n]\!]_{w[\varrho]} \rangle \in [\![R]\!]_w$
- $\mathcal{M}, w[\varrho] \models t_1 \equiv t_2 :\iff [\![t_1]\!]_{w[\varrho]} = [\![t_2]\!]_{w[\varrho]}$
- $\mathcal{M}, w[\varrho] \not\models \bot$
- $\mathcal{M}, w[\varrho] \models \varphi \to \psi :\iff \mathcal{M}, w[\varrho] \not\models \varphi$ or $\mathcal{M}, w[\varrho] \models \psi$
- $\mathcal{M}, w[\varrho] \models \Box_Z \varphi :\iff$ for all $w' \in \mathcal{W}$ such that $w \mathcal{R} w' : \mathcal{M}, w'[Z/w[\varrho]] \models \varphi$

As a corollary to these definitions it can be shown that the semantical interpretation of $\top$, $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \leftrightarrow \psi$ and $\Diamond_Z\varphi$ are as expected:

Corollary 3.13 For any model $\mathcal{M}=\langle\mathcal{W},\mathcal{R}\rangle$, and any $w[\varrho] \in$ WUR:

- $\mathcal{M}, w[\varrho] \models \top$
- $\mathcal{M}, w[\varrho] \models \neg\varphi \iff \mathcal{M}, w[\varrho] \not\models \varphi$
- $\mathcal{M}, w[\varrho] \models \varphi \wedge \psi \iff \mathcal{M}, w[\varrho] \models \varphi$ and $\mathcal{M}, w[\varrho] \models \psi$
- $\mathcal{M}, w[\varrho] \models \varphi \vee \psi \iff \mathcal{M}, w[\varrho] \models \varphi$ or $\mathcal{M}, w[\varrho] \models \psi$
- $\mathcal{M}, w[\varrho] \models \varphi \leftrightarrow \psi \iff \mathcal{M}, w[\varrho] \models \varphi \to \psi$ and $\mathcal{M}, w[\varrho] \models \psi \to \varphi$
- $\mathcal{M}, w[\varrho] \models \Diamond_Z\varphi \iff$ for a $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$: $\mathcal{M}, w'[Z/w[\varrho]] \models \varphi$

Proof: Left to the reader. $\dashv$

The workings of the semantics can be illustrated by the following example, which is a sequel to example 3.8, on page 26, above:

Example 3.14 Consider the m-model $\mathcal{M} = \langle\mathcal{W},\mathcal{R}\rangle$ as it is described in example 3.8. Let $\mathcal{M}$ be further specified as:

- $\mathcal{R} = \mathcal{W} \times \mathcal{W}$
- $[\![t]\!]_{w_1[\text{-}]} = d \neq d' = [\![t]\!]_{w_2[\text{-}]}$
- $[\![P]\!]_{w_1} = \{d\}$
- $[\![P]\!]_{w_2} = \{d'\}$

$\Box_\emptyset Pt$ is true in w_1, as witness the following equivalences:

$\mathcal{M}, w_1[\text{-}] \models \Box_\emptyset Pt$
$\iff$ for all $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$: $\mathcal{M}, w'[\emptyset/w_1[\text{-}]] \models Pt$
$\iff$ $\mathcal{M}, w_1[\emptyset/w_1[\text{-}]] \models Pt$ and $\mathcal{M}, w_2[\emptyset/w_1[\text{-}]] \models Pt$
$\iff$ $[\![t]\!]_{w_1[\emptyset/w_1[\text{-}]]} \in [\![P]\!]_{w_1}$ and $[\![t]\!]_{w_2[\emptyset/w_1[\text{-}]]} \in [\![P]\!]_{w_2}$
$\iff$ $[\![t]\!]_{w_1[\text{-}]} \in [\![P]\!]_{w_1}$ and $[\![t]\!]_{w_2[\text{-}]} \in [\![P]\!]_{w_2}$
$\iff$ $d \in [\![P]\!]_{w_1}$ and $d' \in [\![P]\!]_{w_2}$.

In contrast, $\Box_{\{t\}}Pt$ does not hold in w_1, since $w_1\mathcal{R}w_2$ and:

$[\![t]\!]_{w_2[\{t\}/w_1[\text{-}]]} = [\![t]\!]_{w_1[\text{-}]} = d \notin [\![P]\!]_{w_2}$.

Consequently, $\mathcal{M}, w_2[\{t\}/w_1[\text{-}]] \not\models Pt$ and, finally: $\mathcal{M}, w_1[\text{-}] \not\models \Box_{\{t\}}Pt$

Concerning Frames

In definition 3.7, above, the interpretation function for the non-logical lexicon was taken as an integral part of the worlds in each m-model. Due to this somewhat idiosyncratic definition of models the notion of a frame turns out to be somewhat more complicated than usual. In favour of the definition of models, as it stands, can be said that it neatly preserves the idea that possible worlds are taken to be very much like first-order models and that the notion of frame only plays a marginal role in this thesis.

Since $\mathcal{L}$ is fully interpreted in the worlds of an m-model $\mathcal{M} = \langle\mathcal{W},\mathcal{R}\rangle$, the frame underlying $\mathcal{M}$ cannot be considered to be the ordered pair of $\mathcal{W}$ and $\mathcal{R}$. Therefore, the notion of a frame will be redefined as follows:

Definition 3.15 (Frames) A *frame* $\mathcal{F}$ is a tuple $\langle \mathcal{D}, \mathcal{R} \rangle$, where $\mathcal{D}$ is a set of sets of objects in which the same set of objects may occur several times[4] , and $\mathcal{R} \subseteq \mathcal{D} \times \mathcal{D}$.

Definition 3.16 (Model on a Frame) $\mathcal{M}= \langle \mathcal{W}, \mathcal{R}' \rangle$ is a model on a frame $\mathcal{F}= \langle \mathcal{D}, R \rangle$ iff there exists an isomorphism $f : \mathcal{F} \mapsto \mathcal{M}$ such that for all $D \in \mathcal{D}$: if $f(D) = \langle D'_w, [\![\cdot]\!]_w, g \rangle$ then $D = D'_w$

These definitions preserve many of the characterization results obtained for propositional modal logic.

Concerning Worlds and their Objects

In the semantics above the domains are relative to the worlds. Still, the interpretation function of each wur is total in the sense that the interpretation of each term in a wur ϖ in the domain of ϖ, D_ϖ. By opting for world relative domains, rather then adopting a fixed domain approach, I hoped to secure a maximal degree of flexibility and intuitive plausibility. As matters evolve, however, it turned out that this choice is largely a cosmetic improvement. Every m-model can be demonstrated to be elementary equivalent for some other m-model in which all worlds share the same domain. This is what the following definitions and proofs establish.

Definition 3.17 For any m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ define:

(i) For each $w = \langle D_w, [\![\cdot]\!]_w, g \rangle \in \mathcal{W}$, define:
$w^\ddagger := \langle D_{w^\ddagger}, [\![\cdot]\!]_{w^\ddagger}, g^\ddagger \rangle$, with:

- $D_{w^\ddagger} := \bigcup_{w' \in \mathcal{W}} D_{w'}$
- $[\![\cdot]\!]_{w^\ddagger} := [\![\cdot]\!]_w$
- $g^\ddagger := g$

(ii) $\mathcal{M}^\ddagger := \langle \mathcal{W}^\ddagger, \mathcal{R}^\ddagger \rangle$ with:

- $\mathcal{W}^\ddagger := \{ w^\ddagger \mid w \in \mathcal{W} \}$
- $\mathcal{R}^\ddagger := \{ \langle w^\ddagger, w'^\ddagger \rangle \mid w \mathcal{R} w' \}$

(iii) For each $\varpi \in \text{WUR}_\mathcal{M}$ and $[\varrho] \in \text{REST}_\mathcal{M}$, define $\varpi^\ddagger$ and $[\varrho]^\ddagger$ simultaneously and inductively as:

- $[\text{-}]^\ddagger := [\text{-}]$
- $(w[\text{-}])^\ddagger := w^\ddagger [\text{-}]^\ddagger$
- $[Z/\varpi]^\ddagger := [Z/\varpi^\ddagger]$
- $(w[Z/\varpi])^\ddagger := w^\ddagger [Z/\varpi]^\ddagger$

Note that for each m-model $\mathcal{M}$, $\mathcal{M}^\ddagger$ is also an m-model for $\mathcal{L}$. Observe further that $\{ \varpi^\ddagger \mid \varpi \in \text{WUR}_\mathcal{M} \} = \text{WUR}_{\mathcal{M}^\ddagger}$.
Now it can be shown that each m-model $\mathcal{M}$ is elementary equivalent to $\mathcal{M}^\ddagger$.

Fact 3.18 For all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $w \in \mathcal{W}$:

(i) $[\![t]\!]_\varpi = [\![t]\!]_{\varpi^\ddagger}$

[4] this clause is meant to ensure that there may be models containing different worlds with the same domain.

(ii) $\mathcal{M}, \varpi \models \varphi \Longleftrightarrow \mathcal{M}^{\ddagger}, \varpi^{\ddagger} \models \varphi$

Proof: (i) is by a straightforward induction on both ϖ and t. (ii) is by an equally straightforward induction on φ. $\dashv$

In advance of a notion of validity of formulae in models to be defined below, the following result follows as a corollary.

Corollary 3.19 For each m-model $\mathcal{M}$, $\mathcal{M}$ and $\mathcal{M}^{\ddagger}$ are elementary equivalent.

Proof: Directly from fact 3.18 $\dashv$

If we, however, pledge allegiance to Quine in espousing the view that to be is to be a value of a bound variable rather than the denotation of a singular term, the above results may assume a somewhat spurious aspect. Only if referential devices such as quantifiers or λ-operators are added to the languages of $\mathcal{L}$, there can be sensible talk of domains of quantification. When the quantifiers are added to the language in chapter 6, it will emerge that different choices can be made with respect to the domains of quantification. In opting for world relative domains nothing has been prejudged as to this issue. One may, for instance, decide on $\bigcup_{w \in \mathcal{W}}$, $D_{w[\varrho]}$ or D_w as the domain of quantification of a wur $w[\varrho]$ and different quantified modal systems will ensue. A free modal logic ranks among the possibilities.

Bencivenga characterized free logics as:

> A free logic is a formal system of quantification theory, with or without identity, which allows for some singular terms in some circumstances to be thought of as denoting no existing object, and in which quantifiers are invariably thought of as having existential import. (Bencivenga [1986], p.375)

Understood in this sense, a free interpretation of the quantifiers may be obtained if, for each wur $w[\varrho]$, D_w is taken to be the domain of quantification. If so, it may very well occur that the interpretation of a term in a wur is outside the domain of quantification of that wur. This is a feature that may be welcome to any modal logician or formal semanticist who is interested in possible non-existence of objects and other questions relating to the interaction between existence and modality. Free logics are thought to provide a sounder basis for quantified modal logic than classical predicate logics in this respect (cf. Garson [1984] and Garson [1991]). Whatever the virtues and vices of free logic, it be observed that the notion of truth with respect to formulae containing terms that are undefined in the domain of quantification is far from unequivocal (cf. Bencivenga [1986] and Bencivenga [1991]). With respect to the satisfaction conditions of formulae containing undefined terms widely divergent choices may be feasible with respect to different applications. It is in this sense that the semantical clauses of **PLuM**, as they stand, are not entirely neutral as to the quantification theory with which **PLuM** can be extended. For one thing, if $[\![t]\!]_{w[\varrho]} \notin D_w$ then for each one-place predicate symbol $\mathcal{M}, w[\varrho] \not\models Pt$. Although convenient, this is not an essential feature of all semantics for indexed modal quantifiers. Alternative choices could be made in this respect.

Concerning the Languages of $\mathcal{L}^*$

The semantics being presented, it can also be explained why the restriction to languages in the class $\mathcal{L}^*$ is a natural one. Apart from being a particular convenient stipulation from a technical point of view, it can also be argued for on more philosophical grounds.

If an index is not closed under taking subterms, one is committed to granting intelligibility to formulae in which some composite term $f(t_1, \ldots, t_n)$ gets a *de re* interpretation, whereas one of its subterms should be taken *de dicto*. Yet, intuitively it is not altogether clear what it means if an occurrence of a term t as a subterm of a functional term $f(t_1, \ldots, t_n)$, were to get a *de dicto* interpretation in some modal formula $\Box_Z\varphi(f(t_1, \ldots, t_n))$ with $f(t_1, \ldots, t_n)$ acquiring a *de re* reading. Consider for instance (3.1):

(3.1) Necessarily, the mother of the President of the United States is decent

Suppose someone maintained (3.1) to be true, while explicitly stating that 'the mother of the President of the United States' should be taken *de re* and 'the President of the United States' *de dicto*. What could any such person mean to convey? *De re* necessity concerns the properties an object denoted in the actual world by a term, has in other possible worlds. So, in (3.1) would be about the object referred to in the actual world by 'the mother of the President of the United States' with 'the President of the United States' getting a *de dicto* interpretation. Which respectable lady could this be? Would the reference of 'the mother of the President of the United States' succeed only if everyone who could possibly be President of the United States of the United States has actually the same mother?[5]

Let it be noted that **PLuM** does not provide such *de re* readings of functional terms containing occurrences of subterms that get a *de dicto* interpretation, not even for languages of $\mathcal{L}$ which allow for indices that are not closed under taking subterms. If a functional term $f(t_1, \ldots, t_n)$ occurs within the scope of a modal operator $\Box_Z$ such that $f(t_1, \ldots, t_n) \in Z$, then the occurrences of t as a subterm of $f(t_1, \ldots, t_n)$ within the scope of $\Box_Z$ will also get a *de re* reading, even if $t \notin Z$. Suppose one were to evaluate $\Box_{\{f(t_1,\ldots,t_n)\}}P(f(t_1, \ldots, t_n))$ in some world from which only w' is accessible. One will then end up checking whether $[\![f]\!]_{w_1}([\![t_1]\!]_{w_1}, \ldots, [\![t_n]\!]_{w_1}) \in [\![P]\!]_{w_2}$, just as one would have had, had the index been closed under subterms. When a functional term $f(t_1, \ldots, t_n)$ is in the index of a modal operator it liberates, as it were, its subterms. If $f(t_1, \ldots, t_n)$ occurs within the scope of a modal operator $\Box_Z$ such that $f(t_1, \ldots, t_n) \in Z$, it is immaterial to the interpretation of that occurrence of $f(t_1, \ldots, t_n)$ whether $t \in Z$ or not.

It may be that *de dicto* readings of a subterm t of a functional term $f(t_1, \ldots, t_n)$ that is to get a *de re* interpretation cannot be made good sense of if t occurs as a subterm of $f(t_1, \ldots, t_n)$. But what if a subterm t of $f(t_1, \ldots, t_n)$ has also an autonomous occurrence within the scope of a modal operator $\Box_Z$ with $f(t_1, \ldots, t_n) \in Z$ but $t \notin Z$? Consider for instance:

(3.2) $\Box_{\{\text{the mother of the President of the United States}\}}$ the mother of the President of the United States is older then the President of the United States.

[5] I am indebted to Frank Veltman pointing this out to me.

(3.2) would most naturally be interpreted as expressing a falsehood. The actual mother of the actual President of the United States could surely have run for the office herself and have won the elections. If that had been the case, however, she would certainly not have been older than herself. Yet, had the index

{the mother of the President of the United States}

been closed under taking subterms, the truth-value of the resulting formula could be decided on, only if the matter concerning whether one could have another mother in another world were settled first.

It should be conceded that by restricting one to the languages of $\mathcal{L}^*$ one will remain unable to account for *de re-de dicto* distinctions of this vintage and that taking recourse to languages with indices that are not necessarily closed under taking subterms may provide some relief. Yet, the problem, as I see it, part of a more general shortcoming common to all languages of $\mathcal{L}$. The indices of the languages of $\mathcal{L}$ are such that modal operators indexed with them cannot distinguish between different occurrences of the same term in their scope. Either all autonomous occurrences of a term t in the scope of the same modal operators will obtain a *de re* reading or all of them a *de dicto* interpretation. This may not be entirely satisfactory from a formal semanticist's point of view, who might wish to account for at least three, if not four, different readings of:

(3.3) The President of the United States is necessarily the President of the United States,

This shortcoming of the languages of $\mathcal{L}$ could easily be remedied by indexing the modal operators not just with sets of terms but rather with sequences of sets of terms, $\langle Z_1, \ldots, Z_n \rangle$. Each set in the sequence, Z_i, should then account for the i-th argument position in the subsequent formula as to whether it should get a *de re* or a *de dicto* reading. Note that it would then only be natural to close each of the sets in these sequences under subterms.

One of the reasons why this latter course has not been adopted in this study is that it would needlessly elaborate the notations, whereas the alterations to the logical principles would be only slight. Moreover, it would disturb the symmetry with quantification (cf. chapter 6); any quantifier $\forall x$ binds indiscriminately all free occurrences of x in its scope.

Concerning the Interpretation of Functional Terms

One of the main complicating factors of **PLuM** is that the interpretation of a functional term $f(t_1, \ldots, t_n)$ in a wur, say $w[Z/\varpi]$, may not be dependent of the respective interpretations of f and $t_1, \ldots, t_n$ in $w[Z/\varpi]$ only. If $f(t_1, \ldots, t_n) \in Z$ then it may be the interpretation of f in some entirely different world that is relevant for the interpretation of $f(t_1, \ldots, t_n)$ in $w[Z/\varpi]$. The following fact demonstrates this:

Fact 3.20 Let $f(t_1, \ldots, t_n) \in \textsc{Term}$. Then it is not in general the case, i.e. for each language L of $\mathcal{L}$, that for each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and each $\varpi \in \textsc{wur}$ there a function $\mathbf{h} : (\bigcup_{w \in \mathcal{W}} D_w)^n \rightarrow \bigcup_{w \in \mathcal{W}} D_w$ such that:

$$[\![f(t_1, \ldots, t_n)]\!]_\varpi = \mathbf{h}([\![t_1]\!]_\varpi, \ldots, [\![t_n]\!]_\varpi)$$

Proof: Consider a model $\mathcal{M}= \langle \mathcal{W}, \mathcal{R}\rangle$ such that for some $w, w' \in \mathcal{W}$, some $d_1, d_2 \in D_w \cap D_{w'}$, and some $a, b \in$ CONS such that $a \not\approx b$:

(i) $d_1 \neq d_2$

(ii) $[\![b]\!]_w = d_1$

(iii) $[\![a]\!]_{w'} = d_1$

(iv) $[\![f]\!]_w(d_1) = d_2$

(v) $[\![f]\!]_{w'}(d_1) = d_1$

It just takes an easy check to see that such an m-model does indeed exist. Suppose further, for a *reductio ad absurdum*, that there is an $\mathbf{h} : (\bigcup_{w\in\mathcal{W}} D_w)^n \to \bigcup_{w\in\mathcal{W}} D_w$ for each $w[Z/\varpi] \in \text{WUR}_{\mathcal{M}}$ such that $[\![f(t_1,\ldots,t_n)]\!]_{w[Z/\varpi]} = \mathbf{h}([\![t_1]\!]_{w[Z/\varpi]},\ldots,[\![t_n]\!]_{w[Z/\varpi]})$. Contrary to (i), $d_1 = d_2$ can now be proven. Consider $[\![f(a)]\!]_{w[\{f(a),a\}/w']}$, then the following equations hold:

$d_1 =_{(v)} [\![f]\!]_{w'}(d_1) =_{(iii)} [\![f]\!]_{w'}([\![a]\!]_{w'}) = [\![f(a)]\!]_{w'} = [\![f(a)]\!]_{w[\{f(a),a\}/w']} =_{ass.}$
$\mathbf{h}([\![a]\!]_{w[\{f(a),a\}/w']}) =_{a\in\{f(a),a\}} \mathbf{h}([\![a]\!]_{w'}) =_{(iii)} \mathbf{h}(d_1) =_{(ii)} \mathbf{h}([\![b]\!]_w) =_{b\notin\{f(a),a\}}$
$\mathbf{h}([\![b]\!]_{w[\{f(a),a\}/w']}) =_{ass.} [\![f(b)]\!]_{w[\{f(a),a\}/w']} =_{f(b)\notin\{f(a),a\}} [\![f(b)]\!]_w =$
$[\![f]\!]_w([\![b]\!]_w) =_{(ii)} [\![f]\!]_w(d_1) =_{(iv)} d_2$ ⊣

It could be wondered whether there are any languages $L \in \mathcal{L}$ for which functional terms can be interpreted in the manner indicated above, i. e. whether there are any languages $L \in \mathcal{L}$ such that for all m-models $\mathcal{M}$, all wurs $\varpi \in \text{WUR}^L_{\mathcal{M}}$ and $f(t_1,\ldots,t_n) \in$ TERM there is in general a function $\mathbf{h}$ such that:

$[\![f(t_1,\ldots,t_n)]\!]_\varpi = \mathbf{h}([\![t_1]\!]_\varpi,\ldots,[\![t_n]\!]_\varpi)$.

It appears that such languages do exist, each $L \in \mathcal{L}^{**}$ being a case in point. Recall that the languages L of $\mathcal{L}^{**}$ are those for which $\text{INDEX}_L \subseteq \wp(\text{ATTERM}) \cup \{\text{TERM}\}$.

Proposition 3.21 For all $L \in \mathcal{L}$:

$L \in \mathcal{L}^{**}$

$\Longrightarrow$

for all m-models $\mathcal{M}$, all $\varpi \in \text{WUR}^L_{\mathcal{M}}$ and n-place function symbols ($n > 0$) $f(t_1,\ldots,t_n) \in$ TERM there is a function $\mathbf{h}$ such that:

$[\![f(t_1,\ldots,t_n)]\!]_\varpi = \mathbf{h}([\![t_1]\!]_\varpi,\ldots,[\![t_n]\!]_\varpi)$.

Proof: Consider an arbitrary $L \in L^{**}$. Define for each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$, each $w[\varrho] \in \text{WUR}_{\mathcal{M}}$ and each n-place function symbol f the function $\mathbf{h}^f_{w[\varrho]}$ inductively as:

$$\mathbf{h}^f_{w[\varrho]} = \begin{cases} \mathbf{h}^f_\varpi & \text{if } [\varrho] = [\text{TERM}_L/\varpi] \text{ for some } \varpi \in \text{WUR} \\ [\![f]\!]_w & \text{otherwise} \end{cases}$$

This is defined correctly, since the case in which $\varrho = [\text{-}]$ is also covered. From the definition follows that $\mathbf{h}^f_{w[\text{-}]} = [\![f]\!]_w$. By induction on ϖ can be proven that in general:

$[\![f(t_1,\ldots,t_n)]\!]_\varpi = \mathbf{h}^f_\varpi([\![t_1]\!]_\varpi,\ldots,[\![t_n]\!]_\varpi)$.

So consider an arbitrary m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$ and an equally arbitrary term $f(t_1,\ldots,t_n) \in \text{TERM}_L$:

- $\varpi = w[\text{-}]$. Just observe that:

$[\![f(t_1,\ldots,t_n)]\!]_{w[\text{-}]} = [\![f]\!]_{w[\text{-}]}([\![t_1]\!]_{w[\text{-}]},\ldots,[\![t_n]\!]_{w[\text{-}]}) =$
$\mathbf{h}^f_{w[\text{-}]}([\![t_1]\!]_{w[\text{-}]},\ldots,[\![t_n]\!]_{w[\text{-}]})$.

- $\varpi = w[Z/\varpi']$. Two cases should be distinguished. Either $f(t_1,\ldots,t_n) \in Z$ or $f(t_1,\ldots,t_n) \notin Z$. If the latter the proof is very much like in the previous case. Just observe that on this assumption $Z \neq \text{TERM}_L$, and so:

  $[\![f(t_1,\ldots,t_n)]\!]_{w[Z/\varpi']} = [\![f]\!]_w([\![t_1]\!]_{w[Z/\varpi']},\ldots,[\![t_n]\!]_{w[Z/\varpi']}) =$
  $\mathbf{h}^f_{w[Z/\varpi']}([\![t_1]\!]_{w[Z/\varpi']},\ldots,[\![t_n]\!]_{w[Z/\varpi']})$.

  If the former observe that since $f(t_1,\ldots,t_n) \notin \text{ATTERM}_L$, because $n > 0$, and it being the case that $L \in \mathcal{L}^{**}$, $Z = \text{TERM}_L$. If so, $t_i \in Z$ for $1 \leq i \leq n$. Hence the following equations hold:

  $[\![f(t_1,\ldots,t_n)]\!]_{w[Z/\varpi']} = [\![f(t_1,\ldots,t_n)]\!]_{\varpi'} =_{i.h.}$
  $\mathbf{h}^f_{\varpi'}([\![t_1]\!]_{\varpi'},\ldots,[\![t_n]\!]_{\varpi'}) =_{Z=\text{TERM}_L} \mathbf{h}^f_{w[Z/\varpi']}([\![t_1]\!]_{\varpi'},\ldots,[\![t_n]\!]_{\varpi'}) =_{t_1,\ldots,t_n \in Z}$
  $\mathbf{h}^f_{w[Z/\varpi']}([\![t_1]\!]_{w[Z/\varpi']},\ldots,[\![t_n]\!]_{w[Z/\varpi']})$. $\dashv$

3.1.3 Notions of Validity

In **PLuM** truth in worlds under the empty restriction [-] could be conceived of as the most fundamental notion. Truth of formulae in wurs would then be regarded as a derived concept, merely meant to facilitate the interpretation of modal formulae. Accordingly, it would be most natural to formulate logical validity as pertaining to truth in worlds rather than to truth in wurs in general.

Definition 3.22 (Validity and Logical Consequence) Let in each case $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $\mathcal{C}$ a class of m-models. Define for each language $L \in \mathcal{L}$:

(*i*) $\mathcal{M} \models \varphi :\Longleftrightarrow$ for all $w \in \mathcal{W}$: $\mathcal{M}, w[\text{-}] \models \varphi$

(*ii*) $\Phi \models^{\mathcal{C}} \psi$
$:\Longleftrightarrow$
for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle \in \mathcal{C}$ and all $w \in \mathcal{W}$:
$\mathcal{M}, w[\text{-}] \models \varphi$ for all $\varphi \in \Phi \Longrightarrow \mathcal{M}, w[\text{-}] \models \psi$

N.B.: If $\mathcal{C}$ happens to be the set of all m-models the superscript '$\mathcal{C}$' will usually be omitted. $\emptyset \models \varphi$ is abbreviated to $\models \varphi$

Although the emphasis may be on the possible worlds rather than on the wurs, a notion of universal truth in wurs corresponding to that of validity is also of considerable importance. As a matter of fact, it is exactly with reference to this difference between truth in all worlds and truth in all wurs that the Quine's argument can be seen to fail for **PLuM** (cf. section 6.4 and section 3.2.4, below). In defining a notion of 'validity' for wurs it may be tempting just to quantify over all wurs of the model. It should be noted, however, that only a limited number of wurs is relevant to the interpretation of formulae in a possible world. Moreover, some wurs of a model might not even be relevant with respect to the interpretation of any formula in any world. For instance, in any m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ with $w_1, w_2 \in \mathcal{W}$ and $\langle w_1, w_2 \rangle \notin \mathcal{R}$, for any language $L \in \mathcal{L}$, the wur $w_2[Z/w_1[\text{-}]]$

will not be relevant to the evaluation of any formula $\varphi \in \text{Form}_L$. The following definition specifies which are the relevant wurs for each possible world.[6]

Definition 3.23 ($\textsc{wur}^L_{\mathcal{M}}(w)$, $\textsc{rest}^L_{\mathcal{M}}(w)$, $\textsc{wur}^L_{\mathcal{M}}(\mathcal{W})$) For each language $L \in \mathcal{L}$, each model $\mathcal{M}=\langle\mathcal{W},\mathcal{R}\rangle$, and each $w \in \mathcal{W}$:

(i) define inductively and simultaneously the set of the relevant wurs for w, $\textsc{wur}^L_{\mathcal{M}}(w)$, and the set of relevant restrictions for w, $\textsc{rest}^L_{\mathcal{M}}(w)$, as, respectively, the smallest sets for which:

- $[\text{-}] \in \textsc{rest}^L_{\mathcal{M}}(w)$
- $w[\text{-}] \in \textsc{wur}^L_{\mathcal{M}}(w)$
- $X \subseteq \text{Index}_L$ and $v[\varrho] \in \textsc{wur}^L_{\mathcal{M}}(w)$ and $v\mathcal{R}v'$ $\Longrightarrow$ $v'[X/v[\varrho]] \in \textsc{wur}^L_{\mathcal{M}}(w)$ and $[X/v[\varrho]] \in \textsc{rest}^L_{\mathcal{M}}(v')$

(ii) $\textsc{wur}^L_{\mathcal{M}}(\mathcal{W}) := \bigcup_{w\in\mathcal{W}} \textsc{wur}^L_{\mathcal{M}}(w)$

(iii) $\textsc{rest}^L_{\mathcal{M}}(\mathcal{W}) = \bigcup_{w\in\mathcal{W}} \textsc{rest}^L_{\mathcal{M}}(w)$

N.B.: The sub- and superscripts $\mathcal{M}$ and L will be omitted when no confusion is likely.

Clearly, $\varpi \in \textsc{wur}^L_{\mathcal{M}}(\mathcal{W}) \subseteq \textsc{wur}^L_{\mathcal{M}}$ and $\varpi \in \textsc{rest}^L_{\mathcal{M}}(\mathcal{W}) \subseteq \textsc{rest}^L_{\mathcal{M}}$.

Wurs playing an intermediary role in the process of determining the truth value of modal formulae in possible worlds, it seems only natural to restrict the notions of validity and logical consequence for wurs to $\textsc{wur}^L_{\mathcal{M}}(\mathcal{W})$ in each model. Accordingly, these notions are defined as follows:

Definition 3.24 (Validity and Logical Consequence in Wurs, $\Vdash$) Let in each case $\mathcal{M} = \langle\mathcal{W},\mathcal{R}\rangle$, $w \in \mathcal{W}$ and $\mathcal{C}$ a class of m-models. Define for each language $L \in \mathcal{L}$:

(i) $\mathcal{M}, w \Vdash \varphi :\Longleftrightarrow$ for all $\varpi \in \textsc{wur}(w)$: $\mathcal{M}, \varpi \models \varphi$

(ii) $\mathcal{M} \Vdash \varphi :\Longleftrightarrow$ for all $w \in \mathcal{W}$: $\mathcal{M}, w \Vdash \varphi$
(or alternatively, $\mathcal{M} \Vdash \varphi :\Longleftrightarrow$ for all $\varpi \in \textsc{wur}^L_{\mathcal{M}}(\mathcal{W})$: $\mathcal{M}, \varpi \models \varphi$)

(iii) $\Phi \Vdash^{\mathcal{C}} \psi$
$:\Longleftrightarrow$
for all m-models $\mathcal{M} \in \mathcal{C}$ and all $\varpi \in \textsc{wur}^L_{\mathcal{M}}(\mathcal{W})$:
$\mathcal{M}, \varpi \models \varphi$ for all $\varphi \in \Phi \Longrightarrow \mathcal{M}, \varpi \models \psi$

N.B.: If $\mathcal{C}$ happens to be the set of all m-models the superscript '$\mathcal{C}$' will usually be omitted. $\emptyset \Vdash \varphi$ is usually abbreviated to $\Vdash \varphi$

Proposition 3.25 For all m-models $\mathcal{M} = \langle\mathcal{W},\mathcal{R}\rangle$ and formulae $\varphi \in \text{Form}$:

(i) $\mathcal{M} \Vdash \varphi \Longrightarrow \mathcal{M} \models \varphi$

(ii) $\Vdash \varphi \Longrightarrow \models \varphi$

Proof: Obvious, since $\{w[\text{-}] \mid w \in \mathcal{W}\} \subseteq \textsc{wur}(\mathcal{W})$. $\dashv$

The converses of proposition 3.25, however, do not in general hold, as witness the following example:

[6] Below, when dealing with Kaplan's view on semantics, the notion of a plain wur becomes relevant again, be it with respect to meaning rather than to modality.

Example 3.26 Below (cf. section 3.2.3) it will be demonstrated that $\models a \equiv b \to f(a) \equiv f(b)$. Still, $\not\Vdash a \equiv b \to f(a) \equiv f(b)$, as the following counterexample makes manifest. Consider an m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ with $\mathcal{W}$ containing at least two worlds, $w_1 = \langle D_{w_1}, [\![\cdot]\!]_{w_1}, g \rangle$ and $w_2 = \langle D_{w_2}, [\![\cdot]\!]_{w_2}, h \rangle$ such that $w_1 \mathcal{R} w_2$. Suppose there are at least two objects d_1, d_2 such that $d_1 \in D_{w_1}$ and $d_2 \in D_{w_2}$. Let further the respective interpretation functions of w_1 and w_2 be such that for $a, b \in \text{CONS}_L$ ($a \not\approx b$), and some $f \in \text{FUNC}_L$:

(i) $[\![a]\!]_{w_1} = d_1 = [\![b]\!]_{w_1}$

(ii) $[\![f]\!]_{w_1}(d_1) = d_1$

(iii) $[\![b]\!]_{w_2} = d_1$

(iv) $[\![f]\!]_{w_2}(d_1) = d_2$

Now consider $w_2[\{f(a), a\}/w_1[\text{-}]] \in \text{WUR}(w_1)$, clearly:

$$[\![a]\!]_{w_2[\{f(a),a\}/w_1[\text{-}]]} = [\![a]\!]_{w_1[\text{-}]} = d_1 = [\![b]\!]_{w_2[\text{-}]} = [\![b]\!]_{w_2[\{f(a),a\}/w_1[\text{-}]]}.$$

Still:

$$[\![f(a)]\!]_{w_2[\{f(a),a\}/w_1[\text{-}]]} = [\![f(a)]\!]_{w_1[\text{-}]} = [\![f]\!]_{[\![a]\!]_{w_1[\text{-}]}} = [\![f]\!]_{w_1}(d_1) = d_1 \neq d_2 =$$
$$[\![f]\!]_{w_2}(d_1) = [\![f]\!]_{w_2}([\![b]\!]_{w_2[\text{-}]}) = [\![f]\!]_{w_2}([\![b]\!]_{w_2[\{f(a),a\}/w_1[\text{-}]]}) = [\![f(b)]\!]_{w_2[\{f(a),a\}/w_1[\text{-}]]},$$

and so: $\mathcal{M}, w_2[\{f(a), a\}/w_1[\text{-}]] \not\models a \equiv b \to f(a) \equiv f(b)$.

Yet, these converses do hold for languages $L \in \mathcal{L}^{**}$. The proof of this latter result is a combination of the proofs of fact 3.18 and proposition 3.21 above. The general idea is that in the case of languages in $\mathcal{L}^{**}$, wurs can be conceived of as worlds. This is not in general the case when other languages are concerned. The proof proceeds by constructing for each model $\mathcal{M}$ an elementary equivalent m-model $\mathcal{M}^* = \langle \mathcal{W}^*, \mathcal{R}^* \rangle$, in which for each wur $\varpi \in \text{WUR}_{\mathcal{M}}$ there is a world $w \in \mathcal{W}^*$ that verifies the same formulae.

Definition 3.27 Define for each $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ the model $\mathcal{W}^* = \langle \mathcal{W}^*, \mathcal{R}^* \rangle$ with $\mathcal{W}^*$ and $\mathcal{R}^*$ as follows:

- $\mathcal{W}^* = \{v_\varpi \mid \varpi \in \text{WUR}^{\mathbf{L}^{**}}_{\mathcal{M}}(\mathcal{W})\}$, with each $v_\varpi = \langle D_{v_\varpi}, [\![\cdot]\!]^{\mathcal{M}^*}_{v_\varpi}, g_{v_\varpi} \rangle \in \mathcal{W}^*$ as follows:
  - (i) $D_{v_\varpi} = \bigcup_{w \in \mathcal{W}_{\mathcal{M}}} D_w$
  - (ii) for all $c \in \text{CONS}$: $[\![c]\!]^{\mathcal{M}^*}_{v_\varpi} = [\![c]\!]^{\mathcal{M}}_{\varpi}$
  - (iii) For all n-place function symbols f^n: $[\![f^n]\!]^{\mathcal{M}^*}_{v_\varpi} = \mathbf{h}^f_\varpi$ with $\mathbf{h}^f_\varpi$ defined inductively as in the proof of proposition 3.21, *viz.* for each $w[\varrho] \in \text{WUR}$ as:
  $$\mathbf{h}^f_{w[\varrho]} = \begin{cases} \mathbf{h}^f_\varpi & \text{if } [\varrho] = [\text{TERM}/\varpi] \text{ for some } \varpi \in \text{WUR} \\ [\![f]\!]_w & \text{otherwise} \end{cases}$$
  - (iv) For each n-place relation symbol: $[\![R]\!]^{\mathcal{M}^*}_{v_\varpi} = [\![R]\!]^{\mathcal{M}}_{\varpi}$
  - (v) $g_{v_\varpi}(x) = [\![x]\!]^{\mathcal{M}}_{\varpi}$
- $v_{w[\varrho]} \mathcal{R}^* v_{w'[\varrho']} \iff w \mathcal{R} w'$

N.B.: Since confusion is hardly likely, the superscript $\mathcal{M}$ will be omitted, i.e. $[\![\cdot]\!]^{\mathcal{M}}_w$ will be written as $[\![\cdot]\!]_w$ and $[\![\cdot]\!]^{\mathcal{M}^*}_{v_\varpi}$ as $[\![\cdot]\!]^*_{v_\varpi}$.

Fact 3.28 If $\mathcal{M}$ is an m-model $\mathcal{M}^*$ is an m-model as well.

Proof: Left to the reader to verify. Observe that the definition of the interpretation function $[\![\cdot]\!]^*_{v_\varpi}$ is possible in virtue of our dealing here with languages $L \in \mathcal{L}^{**}$ only. $\dashv$

The following lemma is merely auxiliary to the main result.

Lemma 3.29 For all $L \in \mathcal{L}^{**}$ and all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and all $w[Z/\varpi] \in \text{WUR}^L_{\mathcal{M}}$ and their $*$-counterparts:

(i) for all $t \in \text{TERM}$: $[\![t]\!]^*_{v_{w[Z/\varpi]}} = [\![t]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]}$

(ii) for all $\varphi \in \text{FORM}$: $\mathcal{M}^*, v_{w[Z/\varpi]} \models \varphi \iff \mathcal{M}^*, v_{w[Z/\varpi]}[Z/v_\varpi] \models \varphi$

Proof: Consider an arbitrary $L \in \mathcal{L}^{**}$ and equally arbitrarily choose an m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $w[Z/\varpi] \in \text{WUR}^L_{\mathcal{M}}$.

(i) By induction on t. In each case one should deal separately with $t \in Z$ and $t \notin Z$. As an illustration only the complex case $t \approx f(t_1, \ldots, t_n)$ will be given in full:

- $t \approx f(t_1, \ldots, t_n)$. The case in which $f(t_1, \ldots, t_n) \in Z$ should be distinguished from the one in which $f(t_1, \ldots, t_n) \notin Z$. If the latter the following equations hold:
  $[\![f(t_1, \ldots, t_n)]\!]^*_{v_{w[Z/\varpi]}} = [\![f]\!]^*_{v_{w[Z/\varpi]}}([\![t_1]\!]^*_{v_{w[Z/\varpi]}}, \ldots, [\![t_n]\!]^*_{v_{w[Z/\varpi]}}) =$
  $[\![f]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]}([\![t_1]\!]^*_{v_{w[Z/\varpi]}}, \ldots, [\![t_n]\!]^*_{v_{w[Z/\varpi]}}) =_{i.h}$
  $[\![f]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]}([\![t_1]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]}, \ldots, [\![t_n]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]}) =$
  $[\![f(t_1, \ldots, t_n)]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]}$.
  If, on the other hand, $f(t_1, \ldots, t_n) \in Z$, then $Z = \text{TERM}_L$, since $f(t_1, \ldots, t_n) \notin \text{ATTERM}_L$. Consequently, or each t_i $(1 \leq i \leq n)$, $t_i \in Z$, and so:
  $[\![f(t_1, \ldots, t_n)]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]} =_{f(t_1,\ldots,t_n)\in Z} [\![f(t_1, \ldots, t_n)]\!]^*_{v_\varpi} =$
  $[\![f]\!]^*_{v_\varpi}([\![t_1]\!]^*_{v_\varpi}, \ldots, [\![t_n]\!]^*_{v_\varpi}) = \mathbf{h}^f_\varpi([\![t_1]\!]^*_{v_\varpi}, \ldots, [\![t_n]\!]^*_{v_\varpi}) =_{t_1,\ldots,t_n \in Z}$
  $\mathbf{h}^f_\varpi([\![t_1]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]}, \ldots, [\![t_n]\!]^*_{v_{w[Z/\varpi]}[Z/v_\varpi]}) =_{i.h.}$
  $\mathbf{h}^f_\varpi([\![t_1]\!]_{v_{w[Z/\varpi]}}, \ldots, [\![t_n]\!]_{v_{w[Z/\varpi]}}) =_{Z=\text{TERM}}$
  $\mathbf{h}^f_{w[Z/\varpi]}([\![t_1]\!]_{v_{w[Z/\varpi]}}, \ldots, [\![t_n]\!]_{v_{w[Z/\varpi]}}) =_{def.}$
  $[\![f]\!]_{v_{w[Z/\varpi]}}([\![t_1]\!]_{v_{w[Z/\varpi]}}, \ldots, [\![t_n]\!]_{v_{w[Z/\varpi]}}) = [\![f(t_1, \ldots, t_n)]\!]_{v_{w[Z/\varpi]}}$.

(ii) By an induction on φ. The result can also be recognized to follow immediately from (i) and lemma 3.34(v), below. $\dashv$

Now the following proposition can be demonstrated, with 3.31 as a corollary.

Proposition 3.30 For all $L \in \mathcal{L}^{**}$ and all $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $\varpi \in \text{WUR}^L_{\mathcal{M}}$:

(i) for all $t \in \text{TERM}_L$: $[\![t]\!]_\varpi = [\![t]\!]^*_{v_\varpi}$

(ii) for all $\varphi \in \text{FORM}_L$: $\mathcal{M}, \varpi \models \varphi \iff \mathcal{M}^*, v_\varpi \models \varphi$.

Proof: Consider an arbitrary $L \in \mathcal{L}^{**}$, as well as an arbitrary $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $\varpi \in \text{WUR}^L_{\mathcal{M}}$.

(i) The proof is by induction on t.

- $t \approx c \in \text{CONS}_L$. Trivial, by definition of $[\![\cdot]\!]^*_{v_\varpi}$
- $t \approx x \in \text{VAR}_L$. Just observe that $[\![x]\!]^*_{v_\varpi} = g_{v_\varpi}(x) = [\![x]\!]_\varpi$
- $t \approx f(t_1, \ldots, t_n)$. This case is in turn by induction on ϖ:
  - $\varpi = w[\text{-}]$. Just consider the following equations: $[\![f(t_1, \ldots, t_n)]\!]_{w[\text{-}]} = [\![f]\!]_{w[\text{-}]}([\![t_1]\!]_{w[\text{-}]}, \ldots, [\![t_n]\!]_{w[\text{-}]}) =_{i.h.}$ $\mathbf{h}^f_{w[\text{-}]}([\![t_1]\!]^*_{v_{w[\text{-}]}}, \ldots, [\![t_n]\!]^*_{v_{w[\text{-}]}}) = [\![f]\!]^*_{v_{w[\text{-}]}}([\![t_1]\!]_{v_{w[\text{-}]}}, \ldots, [\![t_n]\!]^*_{v_{w[\text{-}]}}) =$ $[\![f(t_1, \ldots, t_n)]\!]^*_{v_{w[\text{-}]}}$.
  - $\varpi = w[Z/\varpi']$. Observe that there are two induction hypotheses in force. One of these, $i.h.(i)$ is due to the induction on t, the other, $i.h.(ii)$, can be invoked in virtue of the induction on ϖ. The two cases should be distinguished, (a) $Z = \text{TERM}_L$ and (b) $Z \neq \text{TERM}_L$, can be proven by taking recourse to $i.h.(ii)$ and $i.h.(i)$, respectively.

(ii) By induction on the complexity of $\varphi \in \text{FORM}_L$. The atomic cases $\varphi \approx R(t_1, \ldots, t_n)$ and $\varphi \approx t_1 \equiv t_2$ use (i) above. Here only the case $\varphi \approx \Box_Z\psi$ is given in full:

- $\varphi \approx \Box_Z\psi$. The following equivalences hold:

$$
\begin{array}{ll}
\mathcal{M}, \varpi \not\models \Box_Z\psi & \\
\iff & \text{for some } w' \in \mathcal{W} : \ w\mathcal{R}w' \ \& \ \mathcal{M}, w'[Z/\varpi] \not\models \psi \\
\iff_{\text{i.h.}} & \text{for some } w' \in \mathcal{W} : \ w\mathcal{R}w' \ \& \ \mathcal{M}^*, v_{w'[Z/\varpi]} \not\models \psi \\
\iff_{3.29(ii)} & \text{for some } w' \in \mathcal{W} : \ w\mathcal{R}w' \ \& \ \mathcal{M}^*, v_{w'[Z/\varpi]}[Z/v_\varpi] \not\models \psi \\
\iff & \text{for some } v' \in \mathcal{W}^* : \ v_\varpi\mathcal{R}^*v' \ \& \ \mathcal{M}^*, v'[Z/v_\varpi] \not\models \psi \\
\iff & \mathcal{M}^*, v_\varpi \not\models \Box_Z\psi
\end{array}
$$
$\dashv$

Corollary 3.31 For all $L \in \mathcal{L}^{**}$ and all $\varphi \in \text{FORM}_L$:
$\models \varphi \iff \Vdash \varphi$

Proof: Consider an arbitrary $L \in \mathcal{L}^{**}$.
$\Leftarrow$: Cf. proposition 3.25.
$\Rightarrow$: Suppose for some $\varphi \in \text{FORM}_L$, $\not\Vdash \varphi$. Then for some m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and some $w \in \mathcal{W}$ as well as some $[\varrho] \in \text{REST}_\mathcal{M}$: $\mathcal{M}, \varpi \not\Vdash \varphi$. By theorem 3.30, $\mathcal{M}^*, v_\varpi \not\models \varphi$, which allows us to conclude that $\not\models \varphi$. $\dashv$

As a final observation it be noted that the consequence relations $\models$ and $\Vdash$ are classical in the sense of being reflexive, monotone and transitive (the cut-rule holds).

3.2 Semantics Matters

3.2.1 Some Elementary Results

It only takes an easy check to ascertain that $\perp$ and $\rightarrow$ have their usual Boolean interpretations in **PLuM**. The distributivity of $\Box$ over $\rightarrow$ is also inherited from propositional modal logic. The proof is straightforward.

Proposition 3.32 For all $L \in \mathcal{L}$, all $\varphi, \psi \in \text{FORM}_L$, and all $Z \in \text{INDEX}_L$:
$\models \Box_Z(\varphi \to \psi) \to (\Box_Z\varphi \to \Box_Z\psi)$

Proof: Consider an arbitrary $L \in \mathcal{L}$, some $\varphi, \psi \in \text{FORM}_L$ and $Z \in \text{INDEX}_L$. Consider an equally arbitrary m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $w[\varrho] \in \text{WUR}(\mathcal{W})$. Assume that $\mathcal{M}, w[\varrho] \not\models \Box_Z\varphi \to \Box_Z\psi$. Then (1) $\mathcal{M}, w[\varrho] \models \Box_Z\varphi$ and (2) $\models M, w[\varrho] \not\models \Box_Z\psi$. From the latter we know that there is a $w' \in \mathcal{W}$ such that $\mathcal{M}, w'[Z/w[\varrho]] \not\models \psi$. Consider this w' about which, by (1), it can be known that $\mathcal{M}, w'[Z/w[\varrho]] \models \varphi$. Accordingly, $\mathcal{M}, w'[Z/w[\varrho]] \not\models \varphi \to \psi$ and so $\mathcal{M}, w[\varrho] \not\models \Box_Z(\varphi \to \psi)$. ⊣

As a matter of fact, there are at least two ways in which **PLuM** can be seen to be a conservative extension of propositional modal logic. One way is by correlating the propositional variables of propositional modal logic with the atomic formulae of the languages of $\mathcal{L}$. In that case, however, the modal operators should be regarded as consequently indexed with the empty set. Alternatively, propositional modal logic could be conceived of as the subtheory of **PLuM** as restricted to languages that contain only zero-place relation symbols.

The restrictions of any model for any language pertain only to the interpretation of terms. If two restrictions, $[\varrho]$ and $[\varrho']$, are such that in some possible world w the interpretation of any term t under $[\varrho]$ is identical to the interpretation of t in w under $[\varrho']$, $w[\varrho]$ and $w[\varrho']$ are elementary equivalent.

Definition 3.33 ($[\varrho] \sim^w_Z [\varrho']$, $[\varrho] \sim_Z [\varrho']$). For all languages $L \in \mathcal{L}$, all $Z \subseteq \text{TERM}_L$, all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $[\varrho], [\varrho'] \in \text{REST}^L_{\mathcal{M}}$:
(*i*) $[\varrho] \sim^{\mathcal{M},w}_Z [\varrho'] :\Longleftrightarrow$ for all $z \in Z$: $[\![z]\!]_{w[\varrho]} = [\![z]\!]_{w[\varrho']}$

(*ii*) $[\varrho] \sim^{\mathcal{M}}_Z [\varrho'] :\Longleftrightarrow$ for all $w \in \mathcal{W}$: $[\varrho] \sim^{\mathcal{M},w}_Z [\varrho']$

N.B.: When no confusion is likely the superscript $\mathcal{M}$ will be dropped.

$[\varrho] \sim^w_Z [\varrho']$ thus expresses the coincidence of the interpretation of the terms $t \in Z$ in $w[\varrho]$ and $w[\varrho']$. If $[\varrho] \sim_Z [\varrho']$, $w[\varrho]$ and $w[\varrho']$ interpret the terms $t \in Z$ as the same object for any world w. The first part of the following lemma establishes that coincidence of $[\varrho]$ and $[\varrho']$ with respect to the interpretation of all terms is sufficient for $w[\varrho]$ and $w[\varrho']$ verifying the same formulae. The second part states that merely coincidence of the interpretation of the referring terms of a formula is sufficient for φ having the same truth-value in $w[\varrho]$ and $w[\varrho']$.

Lemma 3.34 For all languages $L \in \mathcal{L}$, all $Z \subseteq \text{TERM}_L$, and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $w \in \mathcal{W}$ and $[\varrho], [\varrho'] \in \text{REST}^L_{\mathcal{M}}$:
(*i*) If $[\varrho] \sim^w_{\text{TERM}_L} [\varrho']$ then for all $\varphi \in \text{FORM}_L$: $\mathcal{M}, w[\varrho] \models \varphi \Longleftrightarrow \mathcal{M}, w[\varrho'] \models \varphi$

(*ii*) For all $\varphi \in \text{FORM}_L$: if $[\varrho] \sim^w_{RT(\varphi)} [\varrho']$ then $\mathcal{M}, w[\varrho] \models \varphi \Longleftrightarrow \mathcal{M}, w[\varrho'] \models \varphi$

Proof: Choose an arbitrary language $L \in \mathcal{L}$, and equally arbitrary $Z \subseteq \text{TERM}_L$, $\varphi \in \text{FORM}_L$, m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $w \in \mathcal{W}$ and $[\varrho], [\varrho'] \in \text{REST}^L_{\mathcal{M}}$. Since, obviously, $[\varrho] \sim^w_{\text{TERM}_L} [\varrho']$ entails $[\varrho] \sim^w_{RT(\varphi)} [\varrho']$ for any $\varphi \in \text{FORM}_L$, it will suffice to prove (*ii*).

(*ii*) Assume $[\varrho] \sim^w_{RT(\varphi)} [\varrho']$ The proof is by an induction on the complexity of φ, of which only the induction step $\varphi \approx \Box_Z\psi$ is worth treating in full:

- $\varphi \approx \Box_Z \psi$. Consider the following equivalences:

$$
\begin{array}{ll}
\multicolumn{2}{l}{\mathcal{M}, w[\varrho] \models \Box_Z \psi} \\
\iff & \text{for all } w' \in \mathcal{W} \text{ such that } w\mathcal{R}w' : \ \mathcal{M}, w'[Z/w[\varrho]] \models \psi \\
\iff_{\text{i.h.}} & \text{for all } w' \in \mathcal{W} \text{ such that } w\mathcal{R}w' : \ \mathcal{M}, w'[Z/w[\varrho']] \models \psi \\
\iff & \mathcal{M}, w[\varrho'] \models \Box_Z \psi
\end{array}
$$

To appreciate that the induction hypothesis is applicable in this last case, it suffices to show that:

For all $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$: $[Z/w[\varrho]] \sim^{w'}_{RT(\psi)} [Z/w[\varrho']]$.

The remainder of this proof is devoted to demonstrating this.

Consider an arbitrary $t \in RT(\psi)$. Either $t \in Z$ or $t \notin Z$. If the former, $t \in RT(\Box_Z \psi)$ and so the following equations hold:

$[\![t]\!]_{w'[Z/w[\varrho]]} =_{t \in Z} [\![t]\!]_{w[\varrho]} =_{ass.} [\![t]\!]_{w[\varrho']} =_{t \in Z} [\![t]\!]_{w'[Z/w[\varrho']]}$.

In the latter case an induction on t is in order:

- $t \in$ ATTERM$_L$. Obviously: $[\![t]\!]_{w'[Z/w[\varrho]]} =_{t \notin Z} [\![t]\!]_{w'} =_{t \notin Z} [\![t]\!]_{w'[Z/w[\varrho']]}$.
- $t \approx f(t_1, \ldots, t_n)$. Just consider the following equations:

$[\![f(t_1, \ldots, t_n)]\!]_{w'[Z/w[\varrho]]} =_{f(t_1,\ldots,t_n) \notin Z}$
$[\![f]\!]_{w'}([\![t_1]\!]_{w'[Z/w[\varrho]]}, \ldots, [\![t_n]\!]_{w'[Z/w[\varrho]]}) =_{i.h.}$
$[\![f]\!]_{w'}([\![t_1]\!]_{w'[Z/w[\varrho']]}, \ldots, [\![t_n]\!]_{w'[Z/w[\varrho']]}) =_{f(t_1,\ldots,t_n) \notin Z}$
$[\![f(t_1, \ldots, t_n)]\!]_{w'[Z/w[\varrho']]}$
and we are done. ⊣

The following lemma states some specific cases of restrictions that coincide on the interpretation of certain sets of terms.

Lemma 3.35 For all languages $L \in \mathcal{L}$, $Y, Z \subset$ TERM$_L$, m-models $\mathcal{M}$, $\varpi \in$ WUR:

(*i*) $[\emptyset/\varpi] \sim_Z [\text{-}]$

(*ii*) $[Z/\varpi] \sim_Y [Z \cap Y/\varpi]$

Proof: Consider an arbitrary language $L \in \mathcal{L}$, $Z \subseteq$ TERM$_L$, m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $\varpi \in$ WUR, $w \in \mathcal{W}$ and $t \in$ TERM$_L$.

(*i*) By induction on t.

(*ii*) Consider an arbitrary $t \in Y$. The case in which $t \in Z$ and the one such that $t \notin Z$ should be distinguished. In the latter case, obviously, $t \notin Z \cap Y$ and so we may reason as follows:

$[\![t]\!]_{w[Z/\varpi]} =_{t \notin Z} [\![t]\!]_w =_{t \notin Z \cap Y} [\![t]\!]_{w[Z \cap Y/\varpi]}$.

If $t \in Z$, however, then also $t \in Z \cap Y$ and so:

$[\![t]\!]_{w[Z/\varpi]} =_{t \in Z} [\![t]\!]_\varpi =_{t \in Z \cap Y} [\![t]\!]_{w[Z \cap Y/\varpi]}$. ⊣

Lemma 3.35 facilitates the proof of the following proposition, which conveys that the semantically relevant terms in an index of a modal operator are those that recur as a referring term in the formula that is in the scope of the modal operator.

Proposition 3.36 For all languages $L \in \mathcal{L}$, all $Z \subseteq$ TERM$_L$ and all $\varphi \in$ FORM$_L$:

(i) $\models \Box_Z\varphi \leftrightarrow \Box_{Z\cap RT(\varphi)}\varphi$

(ii) $\models \Box_Z\Box_Y\varphi \leftrightarrow \Box_{Z\cap Y}\Box_Y\varphi$

Proof: Consider an arbitrary language $L \in \mathcal{L}$, and equally arbitrary $Z \subseteq \textsc{Term}_L$ and $\varphi \in \textsc{Form}_L$, m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$, $w \in \mathcal{W}$, $[\varrho] \in \textsc{rest}(w)$.

(i) In virtue of the lemmas 3.35(ii) and 3.34(ii), above, the following equivalences can be appreciated:

$\mathcal{M}, w[\varrho] \models \Box_Z\varphi$
$\iff$ for all $w' \in \mathcal{W}:\ w\mathcal{R}w' \Longrightarrow \mathcal{M}, w'[Z/w[\varrho]] \models \varphi$
$\iff$ for all $w' \in \mathcal{W}:\ w\mathcal{R}w' \Longrightarrow \mathcal{M}, w'[Z \cap RT(\varphi)/w[\varrho]] \models \varphi$
$\iff$ $\mathcal{M}, w[\varrho] \models \Box_{Z\cap RT\varphi}\varphi$.

(ii) Just observe that $RT(\Box_Y\varphi) \subseteq Y$, then almost immediately from (i). $\dashv$

3.2.2 Modal Identity Statements

In **PLuM** the issue concerning contingent identity and contingent identity statements can easily be assessed. Speaking in terms of m-models, it is clear that any object d is identical to itself and to itself only, irrespective of world the domain of which it is an element of. This is obviously equally true if such an object is denoted by some term. Yet, the term that denotes an object in some world may denote an entirely different object in some other possible world. In that sense identity statements are contingent. Accordingly, one would expect that identity statements are *de re* necessary and *de dicto* contingent. The following propositions reflect these insights.

Proposition 3.37 For all $L \in \mathcal{L}$, $Z \in \textsc{Index}_L$, and all $t_1, t_2 \in \textsc{Term}_L$:

(i) $\models t_1 \equiv t_2 \to \Box_Z t_1 \equiv t_2$ $\qquad t_1, t_2 \in Z$

(ii) $\models t_1 \not\equiv t_2 \to \Box_Z t_1 \not\equiv t_2$ $\qquad t_1, t_2 \in Z$

Proof: Consider an arbitrary $L \in \mathcal{L}$ and equally arbitrary $Z \in \textsc{Index}_L$, $t_1, t_2 \in \textsc{Term}$. Assume in both cases that $t_1, t_2 \in Z$. Consider as well an arbitrary m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$, and a $\varpi \in \textsc{wur}$:

(i) Assume that $\mathcal{M}, \varpi \models t_1 \equiv t_2$, i. e. $[\![t_1]\!]_\varpi = [\![t_2]\!]_\varpi$. Let $\varpi = w[\varrho]$. Consider an arbitrary $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$. The following equations hold:

$$[\![t_1]\!]_{w'[Z/\varpi]} =_{t_1\in Z} [\![t]\!]_\varpi =_{ass.} [\![t_2]\!]_\varpi =_{t_2\in Z} [\![t_2]\!]_{w'[Z/\varpi]}$$

(ii) Assume that $\mathcal{M}, \varpi \models t_1 \equiv t_2$, i. e. $[\![t_1]\!]_\varpi \neq [\![t_2]\!]_\varpi$. Let $\varpi = w[\varrho]$. Then it is sufficient to show that for all $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$: $[\![t_1]\!]_{w'[Z/\varpi]} \neq [\![t_2]\!]_{w'[Z/\varpi]}$. We prove something slightly stronger. Consider an arbitrary $w' \in \mathcal{W}$. The following equations hold:

$$[\![t_1]\!]_{w'[Z/\varpi]} =_{t_1\in Z} [\![t]\!]_\varpi \neq_{ass.} [\![t_2]\!]_\varpi =_{t_2\in Z} [\![t_2]\!]_{w'[Z/\varpi]}$$

$\dashv$

Fact 3.38 For all $L \in \mathcal{L}$, $Z \in \textsc{Index}_L$, $t_1, t_2 \in \textsc{Term}_L$:

(i) $\not\models t_1 \equiv t_2 \to \Box_Z t_1 \equiv t_2$ $\qquad \{t_1, t_2\} \not\subseteq Z$

(ii) $\not\models t_1 \not\equiv t_2 \to \Box_Z t_1 \not\equiv t_2$ $\qquad \{t_1, t_2\} \not\subseteq Z$

Proof: For both (i) and (ii), counterexamples can easily be constructed along similar lines as in example 3.26, above. $\dashv$

3.2.3 Substitution

Quine maintained that failure of substitutivity *salva veritate* makes for referential opacity. If a term could not be substituted for a co-referential one in a context, that term did not occur purely referentially. Føllesdal pointed out that one should distinguish between referential and extensional opacity. Referential opacity pertains to the impossibility of quantifiers binding variables within the scope of an intensional context from without. Since **PLuM** contains no referential devices such as quantifiers or λ-operators, the notion of referential opacity of modal operators remains inscrutable. Still, the notion of referential term in a formula, $RT(\varphi)$, is at our disposal. It is with reference to this concept that matters of substitutivity are contemplated in the remainder of this section.

An intuitively acceptable notion of substitution, which, moreover, bears an unmistakable resemblance to substitution in classical predicate logic, can be defined as follows:

Definition 3.39 (Substitution)

- $(t/t')y :\approx \begin{cases} t & \text{if } t' \approx y \\ y & \text{otherwise} \end{cases}$
- $(t/t')c :\approx \begin{cases} t & \text{if } t' \approx c \\ c & \text{otherwise} \end{cases}$
- $(t/t')f^n(t_1, \ldots, t_n) :\approx \begin{cases} t & \text{if } f(t_1, \ldots, t_n) \approx t' \\ f^n((t/t')t_1, \ldots, (t/t')t_n) & \text{otherwise} \end{cases}$
- $(t/t')R^n(t_1, \ldots, t_n) :\approx R^n((t/t')t_1, \ldots, (t/t')t_n)$
- $(t/t')t_1 \equiv t_2 :\approx (t/t')t_1 \equiv (t/t')t_n$
- $(t/t')\bot :\approx \bot$
- $(t/t')(\varphi \to \psi) :\approx (t/t')\varphi \to (t/t')\psi$
- $(t/t')\Box_Z\varphi :\approx \begin{cases} \Box_Z(t/t')\varphi & \text{if } t' \in Z \\ \Box_Z\varphi & \text{otherwise} \end{cases}$

The question that this section will be devoted to is whether a term s' that occurs referentially in a formula φ, can be substituted by any coreferential term s without it precipitating a change of φ's truth-value. It turns out that this is not in general the case. The catch is that if s' is among the referential terms of a modal formula $\Box_Z\psi$, s has to be a member of the index Z as well. If s is not, replacing s' by s wherever it occurs in ψ, s will get bound by $\Box_Z$ and the preservation of truth-value is no longer guaranteed. Still, there are no languages of $\mathcal{L}$ that could secure any index being closed under coreferential terms in all m-models, apart from those that have at most $\emptyset$ and $\textsc{Term}_L$ among their indices. Confining one's attention to models in which substitution is warranted would come down to a restriction to models in which all terms denote the same objects in any possible world. But as Quine has already foreshadowed, a modal apocalypse will be immanent when term-forming devices as strong as $\imath$-operators are added to the language.

The simile of modal operators binding terms in their scope as do quantifiers variables in quantified predicate logic may prove its value here. A notion corresponding to that of a term being free for another in a formula is needed in the modal case. It should be warranted that, if a term t occurs referentially in some formula φ, replacing it by a coreferential term t' does not result in t' getting bound by a modal operator. Matters are slightly complicated by the possibility of t occurring as a subterm of functional terms. Substituting coreferential terms in a formula is feasible only if it does not result in bound functional terms getting liberated or the other way round. The following definition is meant to forestall these possibilities.

Definition 3.40 (s being free for s' in φ). Define inductively:

- s is free for s' in $R(t_1, \ldots, t_n)$
- s is free for s' in $t_1 \equiv t_2$
- s is free for s' in $\perp$
- s is free for s' in $\psi \to \chi$ $:\Longleftrightarrow$ s is free for s' in ψ and s is free for s' in χ
- s is free for s' in $\Box_Z \psi$
  $:\Longleftrightarrow$
  (i) for all $z \in \text{TERM}_L :\ z \in Z \Longleftrightarrow (s/s')z \in Z$, and
  (ii) s is free for s' in ψ

It then turns out that for all languages $L \in \mathcal{L}^*$:

$$\models s \equiv s' \to (\varphi \leftrightarrow (s/s')\varphi). \qquad s \text{ is free for } s' \text{ in } \varphi$$

The proof for this, however, would be by induction on φ. In order to be in position to apply the induction hypothesis some induction loading is in order. Substitution should be proven with reference to wurs rather than mere worlds in order to cover the modal case in which $\varphi \approx \Box_Z \psi$. When applying the semantical definitions, the index of a modal operator recurs in the index of the wur. If so, the binding potential of the index appears not to have diminished at all. Suppose, for instance, that one were to evaluate a formula $\Box_{\{a\}} Pa$ in a world w. Assume further that a and b are such that $a \not\approx b$ but nevertheless coreferential in w, i.e. $[\![a]\!]_w = [\![b]\!]_w$. b is not free for a in $\Box_{\{a\}} Pa$, though b is in Pa. Still, a and b being coreferential in w does not guarantee that a and b are coreferential in $w'[\{a\}/w]$. This may impede a truth-value preserving substitution of b for a. It could be said that the index $\{a\}$ binds the interpretation of b to w' in $w'[\{a\}/w]$. The following definition is meant to capture the restrictions on wurs that correspond to the conditions for a term being free for another in a formula as they are determined by the indices of the modal operators. For want of a better word I call it pervasiveness of a wur for a term with respect to another term:

Definition 3.41 (Pervasiveness of a wur ϖ for s' with respect to a term s.) Define for each $L \in \mathcal{L}$, $Z \in \text{INDEX}_L$, $s, s' \in \text{TERM}_L$, m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $\varpi \in \text{WUR}$, by induction on ϖ:

- $w[\text{-}]$ is pervasive for s' with respect to s

- $w[Z/\varpi']$ is pervasive for s' with respect to s
  $:\Longleftrightarrow$
  (i) for all $z \in \textsc{Term}_L : \ z \in Z \Longleftrightarrow (s/s')z \in Z$, and
  (ii) ϖ' is pervasive for s' with respect to s

The following fact establishes the interrelationship between a term being free for another in a formula and a wur being pervasive for a term with respect to another term.

Fact 3.42 For all $L \in \mathcal{L}$, $Z \in \textsc{Index}_L$, $s, s' \in \textsc{Term}_L$, $\varphi \in \textsc{Form}_L$ and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $w \in \mathcal{W}$, $\varpi \in \textsc{wur}(\mathcal{W})$:
s' is free for s in $\Box_Z\varphi$ and ϖ is pervasive for s' with respect to s
$\Longleftrightarrow$
$w[Z/\varpi]$ is pervasive for s' with respect to s and s' is free for s in φ

Proof: Straightforward and easy. ⊣

With the concept of pervasiveness of a wur with respect to a term at our disposal the following lemma can be proven:

Lemma 3.43 For all $L \in \mathcal{L}^*$, $s, s', t, t' \in \textsc{Term}_L$, $\varphi \in \textsc{Form}_L$ and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $\varpi \in \textsc{wur}$:

(i) $\mathcal{M}, \varpi \models s \equiv s' \to t \equiv (s/s')t$	ϖ is pervasive for s' w.r.t. s
(ii) $\mathcal{M}, \varpi \models s \equiv s' \to (\varphi \leftrightarrow (s/s')\varphi)$	ϖ is pervasive for s' w.r.t. s and s' is free for s in φ

Proof: Consider an arbitrary $L \in \mathcal{L}^*$, and equally arbitrary m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $\varpi \in \textsc{wur}(\mathcal{W})$, $s, s', t \in \textsc{term}_L$ and $\varphi \in \textsc{Form}_L$:

(i) Assume ϖ to be pervasive for s' with respect to s. The proof is then by induction on t.

- $t \in \textsc{Atterm}_L$. Either $s' \approx t$ or $s' \not\approx t$. In either case the proof verges on the trivial. If the former, just observe that $(s/s')t \approx (s/s')s' \approx s$, and so $s \equiv s' \to t \equiv (s/s')t \approx s \equiv s' \to s' \equiv s$, which clearly holds in any wur. If the latter, $s' \not\approx t$ then $(s/s')t \approx t$ and so $s \equiv s' \to t \equiv (s/s')t \approx s \equiv s' \to t \equiv t$.
- $t \approx f(t_1, \ldots, t_n)$. If $s' \approx f(t_1, \ldots, t_n)$ or s' is not a subterm at all of $f(t_1, \ldots, t_n)$ the proof is almost immediate and comparable to the case in which $t \in \textsc{Atterm}_L$. So we may assume that s' be a *proper* subterm of $f(t_1, \ldots, t_n)$. If so, the case is slightly harder but an induction on ϖ will bring home the bacon. Suppose $[\![s']\!]_\varpi = [\![s]\!]_\varpi$
  - $\varpi = w[\text{-}]$. Just consider the following equations:
    $[\![(s/s')f(t_1, \ldots, t_n)]\!]_{w[\text{-}]} =$
    $[\![f((s/s')t_1, \ldots, (s/s')t_n)]\!]_{w[\text{-}]} =_{s' \not\approx f(t_1,\ldots,t_n)}$
    $[\![f]\!]_{w[\text{-}]}([\![(s/s')t_1]\!]_{w[\text{-}]}, \ldots, [\![(s/s')t_n]\!]_{w[\text{-}]}) =_{i.h.}$
    $[\![f]\!]_{w[\text{-}]}([\![t_1]\!]_{w[\text{-}]}, \ldots, [\![t_n]\!]_{w[\text{-}]}) = [\![f(t_1, \ldots, t_n)]\!]_{w[\text{-}]}$.

- $\varpi = w[Z/\varpi']$. Notice that there are now two induction hypotheses in force. The one deriving from the induction on t, $i.h.(i)$, and the one deriving from the induction on ϖ, $i.h.(ii)$. Again two cases should be distinguished. Either $f(t_1, \ldots, t_n) \in Z$ or $f(t_1, \ldots, t_n) \notin Z$. If the former, due to the assumptions that $L \in \mathcal{L}^*$, s' is a proper subterm of $f(t_1, \ldots, t_n)$, we know that $s, s' \in Z$ and so:
  (1) $[\![s']\!]_{\varpi'} =_{s' \in Z} [\![s']\!]_{w[Z/\varpi']} =_{ass.} [\![s]\!]_{w[Z/\varpi']} =_{s \in Z} [\![s]\!]_{\varpi'}$.
  Furthermore, since $w[Z/\varpi']$ is pervasive for s' with respect to s we may take for granted that (2) $(s/s')f(t_1, \ldots, t_n) \in Z$ and that (3) ϖ' is also pervasive for s' with respect to s. Accordingly the following equations are correct:
  $[\![f(t_1, \ldots, t_n)]\!]_{w[Z/\varpi']} =_{f(t_1,\ldots,t_n) \in Z} [\![f(t_1, \ldots, t_n)]\!]_{\varpi'} =_{i.h.(ii)}$
  $[\![(s/s')f(t_1, \ldots, t_n)]\!]_{\varpi'} =_{(2)} [\![(s/s')f(t_1, \ldots, t_n)]\!]_{w[Z/\varpi']}$.
  Note that $i.h.(ii)$ is applicable in virtue of (1) and (3).
  If, however, $f(t_1, \ldots, t_n) \notin Z$ then by $w[Z/\varpi']$ being pervasive for s' with respect to s, we can know (4) $(s/s')f(t_1, \ldots, t_n) \notin Z$. The following equations then hold:
  $[\![f(t_1, \ldots, t_n)]\!]_{w[Z/\varpi']} =_{f(t_1,\ldots,t_n) \notin Z}$
  $[\![f]\!]_{w[Z/\varpi']}([\![t_1]\!]_{w[Z/\varpi']}, \ldots, [\![t_n]\!]_{w[Z/\varpi']}) =_{i.h.(i)}$
  $[\![f]\!]_{w[Z/\varpi']}([\![(s/s')t_1]\!]_{w[Z/\varpi']}, \ldots, [\![(s/s')t_n]\!]_{w[Z/\varpi']}) =_{(4)}$
  $[\![f(t_1, \ldots, t_n)]\!]_{w[Z/\varpi']}$.

(ii) Assume ϖ to be pervasive for s' with respect to s and, moreover, s' to be free for s in φ. Suppose $[\![s]\!]_{\varpi} = [\![s']\!]_{\varpi}$. It is sufficient to prove

$$\mathcal{M}, \varpi \models \varphi \leftrightarrow (s/s')\varphi,$$

which can be achieved by an induction on φ. Only the case $\varphi \approx \Box_Z\psi$ is not entirely straightforward:

- $\varphi \approx \Box_Z\psi$. Let $\varpi = w[\varrho]$. If $s' \notin RT(\Box_Z\psi)$, $(s/s')\Box_Z\psi \approx \Box_Z\psi$, and the proof is trivial. An easy induction on φ can make this manifest. So we may assume that $s' \in RT(\psi) \cap Z$. Since s' is assumed to be free for s in $\Box_Z\psi$, also (1) s' is free for s in ψ and (2) $(s/s')s' \approx s \in Z$. From (2) we can obtain for any $w' \in \mathcal{W}$:
  (3) $[\![s']\!]_{w'[Z/w[\varrho]]} =_{s' \in Z} [\![s']\!]_{w[\varrho]} =_{ass.} [\![s]\!]_{w[\varrho]} =_{(2)} [\![s]\!]_{w'[Z/w[\varrho]]}$.
  Finally, by fact 3.42 it can be known that (4) for any $w' \in \mathcal{W}$ $w'[Z/w[\varrho]]$ is pervasive for s' with respect to s. (1), (2) and (4) make that the induction hypothesis is applicable in the following equivalences:

$$\begin{array}{ll} \mathcal{M}, w[\varrho] \models \Box_Z\psi & \\ \iff & \text{for any } w' \in \mathcal{W}: \ w\mathcal{R}w' \Longrightarrow \mathcal{M}, w'[Z/w[\varrho]] \models \psi \\ \iff_{i.h.} & \text{for any } w' \in \mathcal{W}: \ w\mathcal{R}w' \Longrightarrow \mathcal{M}, w'[Z/w[\varrho]] \models (s/s')\psi \\ \iff & \mathcal{M}, w[\varrho] \models \Box_Z(s/s')\psi \\ \iff_{s' \in Z} & \mathcal{M}, w[\varrho] \models (s/s')\Box_Z\psi \end{array} \qquad \dashv$$

Substitution *salva veritate* follows as a corollary.

Corollary 3.44 For all languages $L \in \mathcal{L}^*$, all $s, s' \in \text{TERM}_L$ and all $\varphi \in \text{FORM}_L$
$\models s \equiv s' \to (\varphi \leftrightarrow (s/s')\varphi)$ $\qquad$ s' free for s in φ

Proof: Directly from lemma 3.43. Just observe that $w[\text{-}]$ is pervasive for any term with respect to any other. $\dashv$

It should be noted that the concept of pervasiveness is a condition on entities of the semantical framework, *viz.* wurs. One could come to wonder whether there is some notion of substitution with respect to which the Leibnizian principle holds in the wurs and that is, moreover, dependent on syntactical features of the formulae only. $[t/\!\!/t']$ is such a notion of substitution, which is not sensitive to the compound structure of functional terms.

Definition 3.45 (Substitution $[t/\!\!/t']$). For all $L \in \mathcal{L}$, $t, t', t'' \in$ TERM and $\varphi \in$ FORM$_L$

- $[t/\!\!/t']t'' :\approx \begin{cases} t & if\ t' \approx t'' \\ t'' & otherwise \end{cases}$
- $[t/\!\!/t']R^n(t_1, \ldots, t_n) :\approx R^n([t/\!\!/t']t_1, \ldots, [t/\!\!/t']t_n)$
- $[t/\!\!/t']t_1 \equiv t_2 :\approx [t/\!\!/t']t_1 \equiv [t/\!\!/t']t_n$
- $[t/\!\!/t']\bot :\approx \bot$
- $[t/\!\!/t'](\varphi \to \psi) :\approx [t/\!\!/t']\varphi \to [t/\!\!/t']\psi$
- $[t/\!\!/t']\Box_Z\varphi :\approx \begin{cases} \Box_Z[t/\!\!/t']\varphi & if\ t' \in Z \\ \Box_Z\varphi & otherwise \end{cases}$

Now the following can be proven:

Proposition 3.46 For all $L \in \mathcal{L}$, for all $s, s', t \in$ TERM$_L$ and $\varphi \in$ FORM$_L$:

(i) $\Vdash s \equiv s' \to t \equiv [s/\!\!/s']t$

(ii) $\Vdash s \equiv s' \to (\varphi \leftrightarrow [s/\!\!/s']\varphi)$ $\qquad s'$ is free for s in φ

Proof: The proof for (i) is highly trivial. If $t \not\approx s'$, $s \equiv s' \to t \equiv t$ should be demonstrated. If on the other hand $t \approx s'$, the consequent is identical to the antecedent. Yet this triviality renders the notion of pervasiveness redundant in the proof for (ii). The proof is then an easy variation on the proof of lemma 3.43(ii). $\dashv$

3.2.4 Replacement

In propositional modal logic replacement of logically equivalent formulae is a universal property of normal systems. In **PLuM**, however, a weak and a strong variant of replacement can be distinguished. Let $[\varphi/\varphi']\psi$ be any sentence that results from ψ by replacing zero or more occurrences of φ' in ψ by φ.

(i) $\Vdash \varphi \leftrightarrow \varphi' \Longrightarrow \models \psi \leftrightarrow [\varphi/\varphi']\psi$ (weak replacement)

(ii) $\models \varphi \leftrightarrow \varphi' \Longrightarrow \models \psi \leftrightarrow [\varphi/\varphi']\psi$ (strong replacement)

Weak replacement can be seen to hold without much ado.

Proposition 3.47 For all $L \in \mathcal{L}$ and all $\varphi, \varphi', \psi \in$ FORM$_L$:

(i) $\Vdash \varphi \leftrightarrow \varphi' \Longrightarrow \Vdash \psi \leftrightarrow [\varphi/\varphi']\psi$

(ii) $\Vdash \varphi \leftrightarrow \varphi' \Longrightarrow \models \psi \leftrightarrow [\varphi/\varphi']\psi$

Proof: Both proofs are straightforward. In both cases if $\varphi' \approx \psi$ the proposition holds immediately. If, on the other hand, $\varphi' \not\approx \psi$ the proposition can be established by induction on ψ. Just observe that, in order to be justified in using the induction hypothesis in the case $\psi \approx \Box_Z\chi$, one needs $\Vdash \varphi \leftrightarrow \varphi'$ and not merely $\models \varphi \leftrightarrow \varphi'$.$\dashv$

In virtue of corollary 3.31, strong replacement follows immediately from proposition 3.47 for all languages $L \in \mathcal{L}^{**}$.

Corollary 3.48 For all languages $L \in \mathcal{L}^{**}$ and all $\varphi, \varphi', \psi \in \mathrm{FORM}_L$:
$\models \varphi \leftrightarrow \varphi' \Longrightarrow \models [\varphi/\varphi']\psi$

Proof: Immediately from proposition 3.47 and corollary 3.31, above. $\dashv$

For other languages in $\mathcal{L}$ something slightly weaker than strong replacement, though stronger than weak replacement, holds. A wur could be assigned a depth, depending on the number of worlds and indices it involves. The depth of a wur is related to the stage the evaluation of a modal formula is in. A world under the empty restriction [-] has a depth of zero. When evaluating a modal formula $\Box_Z\varphi$ in a wur ϖ of depth n, one has to investigate all accessible wurs $w'[Z/\varpi]$. The latter wur is then apportioned a depth of $n+1$.

Definition 3.49 (Depth of a wur). Define for each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and each $\varpi \in \mathrm{WUR}_{\mathcal{M}}$, the depth of ϖ, $\Delta(\varpi)$, inductively as:

- $\Delta(w[\text{-}]) := 0$
- $\Delta(w[Z/\varpi]) := \Delta(\varpi) + 1$

In a similar fashion a modal formulae can be granted a modal depth, depending on the number of modal operators it contains:

Definition 3.50 (Modal depth of a formula, $\Delta(\varphi)$). Define for each $L \in \mathcal{L}$ the modal depth of a formula $\Delta(\varphi)$ inductively as:

- $\Delta(R(t_1, \ldots, t_n)) := 0$
- $\Delta(t_1 \equiv t_2) := 0$
- $\Delta(\bot) := 0$
- $\Delta(\varphi \to \psi) := \max(\Delta(\varphi), \Delta(\psi))$
- $\Delta(\Box_Z\varphi) := \Delta(\varphi) + 1$

A correlation exists between the depth of wurs and the modal depth of formula in which equivalent formulae can be replaced. If two formulae are equivalent in all wurs of a certain depth at most n, they are in general replaceable in all formulae ψ with a modal depth of n or less without loss of truth-value in the worlds under the empty restriction. This also reveals that, with respect to replacement, not so much as a gulf exists between $\Vdash$- and $\models$-equivalence with nothing in between. Rather, there is a gradual transition from $\models$-validity to $\Vdash$-validity, comprising appropriate intermediate notions of validity for wurs of each depth n.

Proposition 3.51 For all $L \in \mathcal{L}$, all $\varphi, \varphi' \in \mathrm{FORM}_L$ and all $n \in \omega$:

for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, and all $\varpi \in \text{WUR}(\mathcal{W})$ such that $\Delta(\varpi) \leq n$: $\mathcal{M}, \varpi \models \varphi \leftrightarrow \varphi'$
$\Longrightarrow$
for all $\psi \in \text{Form}_L$ and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, and $\varpi \in \text{WUR}(\mathcal{W})$ such that $\Delta(\psi) + \Delta(\varpi) \leq n$: $\mathcal{M}, \varpi \models \psi \leftrightarrow [\varphi/\varphi']\psi$

Proof: Consider an arbitrary $L \in \mathcal{L}$, and equally arbitrary $\varphi, \varphi' \in \text{Form}_L$ and $n \in \omega$. Assume that for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and all $\varpi \in \text{WUR}(\mathcal{W})$ such that $\Delta(\varpi) \leq n$: $\mathcal{M}, \varpi \models \varphi \leftrightarrow \varphi'$. We then prove by induction on ψ that for all $\psi \in \text{Form}_L$ and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and all ϖ such that $\Delta(\psi) + \Delta(\varpi) \leq n$:
$\mathcal{M}, \varpi \models \psi \leftrightarrow [\varphi/\varphi']\psi$.
The atomic cases are all analogous, and $\psi \approx \chi_1 \rightarrow \chi_2$ the induction hypothesis can straightforwardly be applied. Accordingly only the cases $\psi \approx R(t_1, \ldots, t_n)$ and $\psi \approx \Box_Z\chi$ are here treated in full:

- $\varphi \approx R(t_1, \ldots, t_n)$. $\Delta(R(t_1, \ldots, t_n)) = 0$. Either $R(t_1, \ldots, t_n) \approx \varphi'$ or $R(t_1, \ldots, t_n) \not\approx \varphi'$. It now remains to be proven that for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $\varpi \in \text{WUR}(\mathcal{W})$ such that $0 + \Delta(\varpi) \leq n$, either $\mathcal{M}, \varpi \models \varphi \leftrightarrow \varphi'$, which had been assumed, or, respectively, $\mathcal{M}, \varpi \models R(t_1, \ldots, t_n) \leftrightarrow R(t_1, \ldots, t_n)$, which goes almost without saying.
- $\psi \approx \Box_Z\chi$. Suppose for some $m \in \omega$, $\Delta(\chi) = m$, and so $\Delta(\Box_Z\chi) = m + 1$. In case $n \leq m$, there are no m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $\varpi \in \text{WUR}(\mathcal{W})$ such that $\Delta(\Box_Z\chi) + \Delta(\varpi) \leq n$ and what needs to be proven holds vacuously. So assume $n \geq m + 1$, in which case $\Delta(\varpi) \leq n - 1$. Let $\varpi = w[\varrho]$ and consider the following equivalences:
  $\mathcal{M}, w[\varrho] \models \Box_Z\chi$
  $\Longleftrightarrow$ for all $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$: $\mathcal{M}, w'[Z/w[\varrho]] \models \chi$
  $\Longleftrightarrow_{\text{i.h.}}$ for all $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$: $\mathcal{M}, w'[Z/w[\varrho]] \models [\varphi/\varphi']\chi$
  $\Longleftrightarrow$ $\mathcal{M}, w[\varrho] \models \Box_Z[\varphi/\varphi']\chi (\approx [\varphi/\varphi']\Box_Z\chi)$.
  Note that for each $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$, $\Delta(w'[Z/w[\varrho]]) \leq n$ and so for each of them: $\mathcal{M}, w'[Z/w[\varrho]] \models \varphi \leftrightarrow \varphi'$ by assumption. Consequently the induction hypothesis was applicable, and we are done. ⊣

As a special case of proposition 3.51 is then obtained:

Corollary 3.52 For all $L \in \mathcal{L}$, all $\varphi, \varphi', \psi \in \text{Form}_L$ and all $n \in \omega$:
for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, and all $\varpi \in \text{WUR}(\mathcal{W})$ such that $\Delta(\varpi) \leq n$: $\mathcal{M}, \varpi \models \varphi \leftrightarrow \varphi'$
$\Longrightarrow$
for all ψ such that $\Delta(\psi) \leq n$: $\models \psi \leftrightarrow [\varphi/\varphi']\psi$

Proof: This is a special case of proposition 3.51. Just note that for $\models \psi \leftrightarrow [\varphi/\varphi']\psi$, $\psi \leftrightarrow [\varphi/\varphi']\psi$ has to hold in all m-models and all wurs $\varpi \in \text{WUR}(\mathcal{W})$ such that $\Delta(\varpi) = 0$ only. ⊣

Strong replacement as such, however, does not for all languages in general. This can readily be appreciated by observing that:

(3.4) $\models (a \equiv b \rightarrow f(a) \equiv f(b)) \leftrightarrow \top.$

Equally easy to prove is that for *any* $Z_1, \ldots, Z_n \in \textsc{Index}_L$:

(3.5) $\models \Box_{Z_1} \ldots \Box_{Z_n} (\top \leftrightarrow \top)$.

Example 3.26 on page 37 illustrated how:

(3.6) $\mathrel{|\!\not\models} a \equiv b \rightarrow f(a) \equiv f(b)$,

from which follows:

(3.7) $\mathrel{|\!\not\models} (a \equiv b \rightarrow f(a) \equiv f(b)) \leftrightarrow \top$.

And so, for some $Z_1, \ldots, Z_n \in \textsc{Index}_L$:

(3.8) $\not\models \Box_{Z_1} \ldots \Box_{Z_n} ((a \equiv b \rightarrow f(a) \equiv f(b)) \leftrightarrow \top)$.

Accordingly, replacing the first occurrence of $\top$ in (3.5) by $a \equiv b \rightarrow f(a) \equiv f(b)$ does not in general result in the formula retaining its truth-value.

Philosophical questions could be raised as to the significance of this failure of strong replacement in the general case. Equivalence of formulae is usually taken to be the criterion for semantical identity. It could be wondered, however, what remains of semantical identity if replacement of semantically identical formulae is not guaranteed to preserve truth-value. The moral that should be drawn, however, is, not that the concept of equivalence has been compromised in **PLuM**, but sooner that semantical identity pertains to material equivalence in all relevant wurs rather than to sameness of truth value in all possible worlds only. Kaplan's views on the semantics of directly referential expressions, such as anaphora, proper names, demonstratives and other indexicals, sustain this view, as will be argued for in section 5.2. Moreover, the failure of strong substitutivity turns out to be exactly the feature of **PLuM** that makes that Quine's conclusion as to the collapse of modal distinctions in quantified modal logic can be avoided (cf. section 6.4).

3.2.5 Necessitation

There are two notions of validity operative, $\models$ and $\mathrel{|\!\models}$. Moreover numerous indices can be attached to the modal operators in the languages of $\mathcal{L}$. As a result there are quite a few variants of the necessitation schema, not all of which are valid in **PLuM**. Two variants of necessitation deserve special attention:

(i) for all $Z \in \textsc{Index}_L$: $\mathrel{|\!\models} \varphi \Longrightarrow \models \Box_Z \varphi$

(ii) for all $Z \in \textsc{Index}_L$: $\models \varphi \Longrightarrow \models \Box_Z \varphi$

Since $\mathrel{|\!\models} \varphi$ entails $\models \varphi$, (i) will be referred to as *weak necessitation* and the latter as *strong necessitation.* It turns out that weak necessitation holds no matter what language $L \in \mathcal{L}$ we are dealing with. To appreciate this, it is sufficient to prove an even weaker notion of necessitation:

Proposition 3.53 For all $L \in \mathcal{L}$, and for all m-models $\mathcal{M}$:

(i) $\mathcal{M} \mathrel{|\!\models} \varphi \Longrightarrow \mathcal{M} \mathrel{|\!\models} \Box_Z \varphi$

(ii) $\mathrel{|\!\models} \varphi \Longrightarrow \mathrel{|\!\models} \Box_Z \varphi$

Proof: Consider an arbitrary $L \in \mathcal{L}$ and an arbitrary $Z \in \textsc{Index}_L$.

(i) Suppose for an arbitrary m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ that $\mathcal{M} \mathrel{|\!\models} \varphi$. Consider also an arbitrary $w[\varrho] \in \textsc{wur}(\mathcal{W})$, and an arbitrary $w' \in \mathcal{W}$ such that $w \mathcal{R} w'$. Clearly $w'[Z/w[\varrho]] \in \textsc{wur}(\mathcal{W})$. By the assumption it then holds that $w'[Z/w[\varrho]] \models \varphi$, and with w' and $w[\varrho]$ having both been chosen arbitrarily, we are done.

(ii) Directly from (i). $\dashv$

Weak necessitation follows as a corollary of proposition 3.53:

Corollary 3.54 For all $L \in \mathcal{L}$ and all $\varphi \in \textsc{Form}_L$:
$\models \varphi \Longrightarrow \models \Box_Z \varphi$

Proof: Recall proposition 3.25(ii). $\dashv$

Strong necessitation is harder to prove. As a matter of fact, it does not hold in general for all languages $L \in \mathcal{L}$. For instance, $\models a \equiv b \rightarrow f(a) \equiv f(b)$; still $\not\models \Box_{\{f(a),a\}}(a \equiv b \rightarrow f(a) \equiv f(b))$ is an immediate result from example 3.26 on page 37, above.

Strong necessity can be seen to hold for all languages $L \in \mathcal{L}^{**}$.

Fact 3.55 For all languages $L \in \mathcal{L}^{**}$ and for all $Z \in \textsc{Index}_L$:
$\models \varphi \Longrightarrow \models \Box_Z \varphi$

Proof: In virtue of the corollaries 3.54 and 3.31. $\dashv$

With the hopes of establishing the general case of strong necessitation — inferring $\models \Box_Z \varphi$ from $\models \varphi$ for *all* indices Z for all languages $L \in \mathcal{L}$ — being shattered, it may be wondered for which indices it remains a sound principle of inference. In this respect the following result can easily be obtained:

Proposition 3.56 For all $L \in \mathcal{L}$ and all $\varphi \in \textsc{Form}_L$:

(i) $\mathcal{M} \models \varphi \Longrightarrow \mathcal{M} \models \Box_\emptyset \varphi$

(ii) $\models \varphi \Longrightarrow \models \Box_\emptyset \varphi$

Proof: Consider an arbitrary $L \in \mathcal{L}$ and equally arbitrary $\varphi \in \textsc{Form}_L$:

(i) Suppose that for some m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and some $w \in \mathcal{W}$: $\mathcal{M}, w \not\models \Box_\emptyset \varphi$. Then there is a $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$ and $\mathcal{M}, w'[\emptyset/w] \not\models \varphi$. Clearly $[\emptyset/w] \sim_{RT(\varphi)} [\text{-}]$, and so by lemma 3.34 we may conclude that $\mathcal{M}, w' \not\models \varphi$ and subsequently $\mathcal{M} \not\models \varphi$.

(ii) Immediately from (i). $\dashv$

The main reason for fact 3.55 not holding for all languages $L \in \mathcal{L}$ was that it is not in general the case m-models can be redefined in such a manner that for each wur there is an elementary equivalent world. It appears, however, that for each φ and m-model $\mathcal{M}$ such that φ holds in a certain kind of wur in $\mathcal{M}$, there is an m-model $\mathcal{M}'$ such that φ holds in a world of $\mathcal{M}'$. The actual proposition and the accompanying proof may strike one as slightly complicated and perhaps a little obscure. Also some preliminaries are in order.

Definition 3.57 Define for each $L \in \mathcal{L}$:

- $\{\breve{x}\} := \{t \in \textsc{Term}_L \mid x \in RT(t)\}$
- $\{\breve{c}\} := \{t \in \textsc{Term}_L \mid c \in RT(t)\}$

- $\{\breve{f}\} := \{t \in \text{TERM}_L \,|\, \text{for some } t_1, \ldots, t_n \in \text{TERM}_L : \ f(t_1, \ldots, t_n) \in RT(t)\}$

Then define for each $F \subseteq \text{FUNC}_L \cup \text{VAR}_L$: $\breve{F} := \bigcup_{f \in F}\{\breve{f}\}$.
Let further $\breve{\bar{F}} := \text{TERM}_L \backslash \breve{F}$

With respect to the soundness proof in the next chapter, it is important to note that for each $F \subseteq \text{FUNC}_L \cup \text{VAR}_L$, $\breve{\bar{F}} \in \text{INDEX}_{\mathbf{L}^*}$. A more general necessitation rule than in proposition 3.56 may now be obtained for a considerable class of languages, of which $\mathbf{L}^*$ is a member.

Proposition 3.58 For all $L \in \mathcal{L}$, and all $\varphi \in \text{FORM}_L$, $F \subseteq \text{FUNC}_L \cup \text{VAR}_L$ such that $\breve{\bar{F}} \in \text{INDEX}_L$. If for all $t \in RT(\varphi) \cap \breve{F} : t \approx f(t_1, \ldots, t_n)$ for some $f \in F$, then:
$\models \varphi \Longrightarrow \models \Box_{\breve{\bar{F}}}\varphi$

Proof: (sketch) Consider an arbitrary $L \in \mathcal{L}$, and equally arbitrary $\varphi \in \text{FORM}_L$, $F \subseteq \text{FUNC}_L \cup \text{VAR}_L$ such that $\breve{\bar{F}} \in \text{INDEX}_L$. Assume that $\not\models \Box_{\breve{\bar{F}}}\varphi$. Consequently, there is an m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and a $w \in \mathcal{W}$ such that $\mathcal{M}, w[\text{-}] \not\models \Box_{\breve{\bar{F}}}\varphi$. We now may assume the existence of a $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$ and $\mathcal{M}, w'[\breve{\bar{F}}/w[\text{-}]] \not\models \varphi$. The idea is now to construct an m-model which contains a world, rather than a wur, that falsifies φ. For this purpose define $\mathcal{M}^\dagger = \langle \mathcal{W}^\dagger, \mathcal{R}^\dagger \rangle$ as follows:

- $\mathcal{W}^\dagger := \mathcal{W} \cup \{v_{w'[\breve{\bar{F}}/w[\text{-}]]}\}$ with
  $v_{w'[\breve{\bar{F}}/w[\text{-}]]} = \langle D_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}}, [\![\cdot]\!]_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}}, g_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}} \rangle$ such that:
  - (i) $D_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}} = D_{w'[\breve{\bar{F}}/w[\text{-}]]}$
  - (ii) for all $c \in \text{CONS}_L$: $[\![c]\!]_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}} := [\![c]\!]_{\mathcal{M}, w'[\breve{\bar{F}}/w[\text{-}]]}$
  - (iii) for all $R \in \text{REL}_L$: $[\![R]\!]_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}} := [\![R]\!]_{w'}$
  - (iv) for all $f \in \text{FUNC}_L$:
    $$[\![f]\!]_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}} := \begin{cases} [\![f]\!]_{w'} & \text{if } f \in F \\ [\![f]\!]_{w} & \text{otherwise} \end{cases}$$
  - (v) for all $x \in \text{VAR}_L$: $g_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}}(x) := [\![x]\!]_{\mathcal{M}, w'[\breve{\bar{F}}/w[\text{-}]]}$
- $\mathcal{R}^\dagger := \mathcal{R} \cup \{\langle v_{w'[\breve{\bar{F}}/w[\text{-}]]}, w'' \rangle \,|\, w'\mathcal{R}w''\}$

For induction loading purposes we now prove:

(1) For all $t \in \text{TERM}_L$ such that either $t \in \breve{\bar{F}}$ or $t \approx f(t_1, \ldots, t_n)$ for some $f \in F$, and for all $w_1, \ldots, w_n \in \mathcal{W}$ and $Z_1, \ldots, Z_n \in \text{INDEX}_L$:
$[\![t]\!]_{w_n[Z_n/\ldots/w_1[Z_1/w'[\breve{\bar{F}}/w[\text{-}]]\ldots]} = [\![t]\!]^\dagger_{w_n[Z_n/\ldots/w_1[Z_1/v_{w'[\breve{\bar{F}}/w[\text{-}]]}]\ldots]}$

The proof of this is by induction on $\Delta(w_n[Z_n/\ldots/w_1[Z_1/w'[\breve{\bar{F}}/w[\text{-}]]\ldots]) \geq 1$ (cf. definition 3.49, page 48). The basic step — for all $t \in \text{TERM}_L$ such that either $t \in \breve{\bar{F}}$ or $t \approx f(t_1, \ldots, t_n)$ for some $f \in F$: $[\![t]\!]_{w'[\breve{\bar{F}}/w[\text{-}]]} = [\![t]\!]^\dagger_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}}$ — is then in turn by induction on t, of which here only the induction step $t \approx f(t_1, \ldots, t_n)$ will be treated in full:

- $t \approx f(t_1, \ldots, t_n)$. Two cases should be distinguished: either $f(t_1, \ldots, t_n) \in \breve{\bar{F}}$ or $f(t_1, \ldots, t_n) \notin \breve{\bar{F}}$. If the former observe that then $f \notin F$ and for each i: $(1 \leq i \leq n)$, $t_i \in \breve{\bar{F}}$, and so the following equations hold:
  $[\![f(t_1, \ldots, t_n)]\!]_{w'[\breve{\bar{F}}/w[\text{-}]]} = [\![f(t_1, \ldots, t_n)]\!]_{w[\text{-}]} =$
  $[\![f]\!]_w([\![t_1]\!]_{w[\text{-}]}, \ldots, [\![t_n]\!]_{w[\text{-}]}) =_{i.h.} [\![f]\!]_w([\![t_1]\!]^\dagger_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}}, \ldots, [\![t_n]\!]^\dagger_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}}) =_{f \notin F}$
  $[\![f()_1, \ldots,)_n]\!]^\dagger_{v_{w'[\breve{\bar{F}}/w[\text{-}]]}}$.

If the latter, however, $f(t_1, \ldots, t_n) \in \breve{F}$. We may assume that $f \in F$. Accordingly, the following equation now hold:

$[\![f(t_1, \ldots, t_n)]\!]_{w'[\bar{\breve{F}}/w[\text{-}]]} = [\![f]\!]_{w'}([\![t_1]\!]_{w'[\bar{\breve{F}}/w[\text{-}]]}, \ldots, [\![t_n]\!]_{w'[\bar{\breve{F}}/w[\text{-}]]}) =_{i.h.}$
$[\![f]\!]_{v_{w'[\bar{\breve{F}}/w[\text{-}]]}}([\![t_1]\!]^{\dagger}_{v_{w'[\bar{\breve{F}}/w[\text{-}]]}}, \ldots, [\![t_1]\!]^{\dagger}_{v_{w'[\bar{\breve{F}}/w[\text{-}]]}}) = [\![f(t_1, \ldots, t_n)]\!]^{\dagger}_{v_{w'[\bar{\breve{F}}/w[\text{-}]]}}$.

The induction step is easy. Since there is no $w \in \mathcal{W}^{\dagger}$ such that $w\mathcal{R}^{\dagger}v_{w'[\bar{\breve{F}}/w[\text{-}]]}$[7], it can now easily be established that for any $w \in \mathcal{W}$: $\mathcal{M}, w \models \varphi \iff \mathcal{M}^{\dagger}, w \models \varphi$. By induction on φ — the induction hypothesis is only needed in the case that $\varphi \approx \psi \to \chi$ — it can be demonstrated that:

(2) $\mathcal{M}, w'[\bar{\breve{F}}/w[\text{-}]] \models \varphi \iff \mathcal{M}^{\dagger}, v_{w'[\bar{\breve{F}}/w[\text{-}]]} \models \varphi$,

and we are done. ⊣

This necessitation rule may not immediately appeal to one's sense of elegance and austerity. Still, I daresay, it can be vindicated on at least two counts. First, proposition 3.58 also covers the case in which $F = \emptyset$. Hence, the following necessitation rule is obtained for all languages $L \in \mathcal{L}$ with $\text{TERM}_L \in \text{INDEX}_L$:

$\models \varphi \Longrightarrow \models \Box_{\text{TERM}_L}\varphi$.

Second, proposition 3.56 is the special case in which $F = \text{FUNC}_L \cup \text{VAR}_L$. All other cases are neat intermediate cases. The unattractive restrictions were called for to render the result eligible for the languages $L \in \mathcal{L}^*$, in particular for $\mathbf{L}^*$. With respect to this latter language of $\mathcal{L}$ soundness and completeness will be proven in the next chapter. In the course of the adequacy proof one of the intermediate rules will be utilized. Accordingly, the following result can be obtained as a corollary.

Corollary 3.59 For $\mathbf{L}^*$, and all $\varphi \in \text{FORM}_L$:

(*i*) $\models \varphi \Longrightarrow \models \Box_{\text{TERM}_L}\varphi$

(*ii*) $\models \varphi \Longrightarrow \models \Box_{RT(\varphi)}\varphi$

Proof: (*i*) is the trivial case of proposition 3.58 in which $F = \emptyset$. The proof of (*ii*) follows almost immediate from (*i*) and lemma 3.36, above. ⊣

[7]There is a problem here. $\mathcal{R}^{\dagger}$ need not be reflexive or symmetric if $\mathcal{R}$ is. This means that the proof of the necessitation rule as it is given here, is not a valid proof for the **T**, **S4** or **S5** extensions of the system.

Chapter 4

Completeness

In this chapter an axiom system for $\mathbf{L}^*$ (for each lexicon L) will be presented that is both sound and complete with respect to the semantics of **PLuM**.

4.1 Axioms and Soundness

The formula schemes that axiomatize **PLuM** for $\mathbf{L}^*$ comprise all Boolean schemes, distribution of $\Box_Z$ over $\rightarrow$ for every $Z \in \textsc{Index}_{\mathbf{L}^*}$, and three axioms governing identity and substitution, as well as *modus ponens* and a rule of necessitation.

Definition 4.1

(A1) all Boolean tautologies

(A2) $\Box_Z(\varphi \rightarrow \psi) \rightarrow (\Box_Z\varphi \rightarrow \Box_Z\psi)$

(A3) $\Box_Z\varphi \leftrightarrow \Box_{Z \cap RT(\varphi)}\varphi$

(A4) $t \equiv t$

(A5) $s \equiv s' \rightarrow ((s'/t)\varphi \rightarrow (s/t)\varphi)$ $\qquad$ s, s' free for t in φ

(A6) $t_1 \not\equiv t_2 \rightarrow \Box_Z t_1 \not\equiv t_2$ $\qquad$ if $t_1, t_2 \in Z$

(MP) If $\Phi \vdash \varphi \rightarrow \psi$ and $\Phi \vdash \varphi$, then $\Phi \vdash \psi$

(Nec) For $F \subseteq \textsc{Func}_L \cup \textsc{Var}_L$:
if for all $t \in RT(\varphi) \cap \breve{F} : t \approx f(t_1, \ldots, t_n)$ for some $f \in F$, then:
$\vdash \varphi \Longrightarrow \vdash \Box_{\breve{F}}\varphi$

We write $\Phi \vdash \varphi$ if there exists a derivation of φ using only $\psi \in \Phi$ as premises, (A1)–(A6) as axioms and (MP) and (Nec) as the sole rules of inference.

In proving completeness of **PLuM** with respect to $\mathbf{L}^*$ we demonstrate that for all $\Gamma \cup \{\varphi\} \subseteq \textsc{Form}_{\mathbf{L}^*}$:

$$\Gamma \models \varphi \iff \Gamma \vdash \varphi$$

The $\Leftarrow$-part of this claim, the soundness of the axiom system, can be assembled from various theorems and propositions proven in the previous section:

Theorem 4.2 (Soundness) For $\mathbf{L}^*$, and all $\Gamma \cup \{\varphi\} \subseteq \textsc{Form}_{\mathbf{L}^*}$:

$$\Gamma \vdash \varphi \Longrightarrow \Gamma \models \varphi$$

Proof: In section 3.2.1 it has been observed that $\perp$ and $\rightarrow$ get there usual interpretations in **PLuM**. This settles the case for $(A1)$ and (MP). The soundness of (A4) is hardly worth proving. It only be noted that the indices of the languages of $\mathcal{L}$ do not distinguish between different occurrences of the same term within their scope. The soundness remaining axioms have all been proven in the previous section: (A2) by proposition 3.32 on page 40, (A3) by proposition 3.36(i) on page 41, (A5) by corollary 3.44 on page 46 and (A6) by proposition 3.37 on page 42. The apparently somewhat complicated necessitation rule was proven in the last section of the previous chapter through proposition 3.37(ii) on page 52. $\dashv$

The remainder of this chapter will be concerned with the $\Rightarrow$ direction, the adequacy of the axiom system, which is rather harder to prove.

4.2 Adequacy

Proving adequacy of the axiom system is considerably harder and rather more complicated. In order to achieve it, standard techniques as they are used in the Henkin-style adequacy proofs of both first-order quantification theory and propositional modal logic will be applied. The general set-up of the proof will be by constructing a canonical m-model for $\mathbf{L}^*$, which constitutes a counter example to $\Gamma \models \varphi$ whenever $\Gamma \not\vdash \varphi$.

As in the propositional modal case, maximal consistent theories will be resorted to in order to supply the necessary building blocks of the canonical m-model. The maximal consistent theories will each determine the interpretation function of a possible world. Yet, mere maximal consistent theories in $\mathbf{L}^*$ will not suffice. The language from which the objects of the canonical m-model are to be drawn should be more comprehensive. In **PLuM** the references of terms in a world have also to be accounted for. This can be achieved by taking equivalence classes of terms of the language from which the canonical model is to be built. These equivalence classes are determined by the identity-statements in the maximal consistent theory associated with the world. If the language in which the maximal consistent theories are formulated, were to coincide with the language for which adequacy has to be proven, it would only be all too reasonable to have $[\![t]\!]_w = \{t' \mid t \equiv t' \in \Gamma\}$, where Γ is the maximal consistent theory underlying w. In each world w of the canonical model, all objects in the domain of w would then be denoted by some term. Moreover, the extension of the relation symbols should also be drawn from this domain. Eventually, this would render it impossible to account for the consistency of $\Box_\emptyset Pa \wedge \neg\Box_{\{a\}} Pa$. The problem is reminiscent of the one one may encounter when proving adequacy of first-order quantification theory. In order to guarantee that the maximal consistent theories from which the canonical model was to be built, i. e. Henkin-theories, one could add a countably infinite number of new individual constants to the language, which are to serve as witnesses to each formula $\exists x\varphi$.

A similar solution will be offered here for **PLuM**. The language from which the canonical m-model $\mathcal{M}_{\mathbf{PLuM}}$ will be built will be $\mathbf{L}_\omega$, as defined below:

Definition 4.3 (The language $\mathbf{L}_\omega$) For each lexicon L define L_ω:

(i) The non-logical lexicon L_ω:

- if f is a function symbol of L then for each $\mathbf{n} \in \omega$: $f_\mathbf{n}$ is a function symbol of L_ω
- if R is a relation symbol of L then R is also a relation symbol of L_ω
- if $x \in \mathrm{VAR}_\mathrm{L}$ then $x_\mathbf{n} \in \mathrm{VAR}_{\mathrm{L}_\omega}$

(ii) $\mathrm{INDEX}_{\mathbf{L}_\omega} := \{Z \subseteq \mathrm{TERM}_{\mathbf{L}_\omega} \mid Z$ is closed under subterms$\}$
(i. e. $\mathbf{L}_\omega = \mathbf{L}^*_\omega$. Obviously, $\mathbf{L}_\omega \in \mathcal{L}^*$.)

For each maximal consistent theory $\Gamma \in \mathrm{FORM}_{\mathbf{L}_\omega}$ and each boldface integer $\mathbf{n}$, $\Gamma_\mathbf{n}$ will then be a world of the canonical m-model. A set of equivalence classes of terms $\mathbf{t} \in \mathrm{TERM}_{\mathbf{L}_\omega}$ will constitute the domains of any such world, in which each $t \in \mathrm{TERM}_{\mathbf{L}^*}$ is going to denote one of these equivalence classes. This does not prevent there being equivalence classes of terms of $\mathbf{L}_\omega$ in the domain of some world $\Gamma_\mathbf{n}$ that are *not* denoted by some term $t \in \mathbf{L}^*$ in $\Gamma_\mathbf{n}$.

In the propositional case the maximal consistent theories could conveniently be identified with the worlds of the canonical model $\mathcal{M}$ and adequacy could be clinched by proving that:

$$\mathcal{M}, \Gamma \models \varphi \iff \varphi \in \Gamma$$

For **PLuM** this would not do on two counts. First, the language in which the maximal consistent are formulated is disjoint from the one for which adequacy should be proven. Consequently, at least some translation is in order. Secondly, not only truth in worlds but also truth in all relevant wurs should be settled. This can be accomplished by choosing the translation function cleverly. Let $\mathcal{M}_\mathbf{PLuM} = \langle \mathcal{W}, \mathcal{R} \rangle$ be the canonical m-model for **PLuM**. What we are then looking for is some translation $*$ from $\mathbf{L}_\omega$ into $\mathbf{L}^*$ such that for each $\Gamma_\mathbf{n}[\varrho] \in \mathrm{WUR}(\mathcal{W})$:

$$\mathcal{M}_\mathbf{PLuM}, \Gamma_\mathbf{n}[\varrho] \models \varphi \iff \varphi^* \in \Gamma$$

Since an intimate relation obtains between $\Gamma_\mathbf{n}[\varrho]$ and the translation asked for, the wurs of the canonical m-model double as translation functions from $\mathbf{L}_\omega$ into $\mathbf{L}^*$:

Definition 4.4 ($\Gamma_\mathbf{n}[\varrho]$-translation) For each wur $\Gamma_\mathbf{n}[\varrho]$ of the canonical m-model $\mathcal{M}_\mathbf{PLuM}$ define:

(i) for all $t \in \mathrm{TERM}_{\mathbf{L}^*}$:

- $c^{\Gamma_\mathbf{n}[\varrho]} :\approx \begin{cases} c^\varpi & \text{if } [\varrho] = [Z/\varpi] \text{ and } c \in Z \\ c_\mathbf{n} & \text{otherwise} \end{cases}$
- $x^{\Gamma_\mathbf{n}[\varrho]} :\approx \begin{cases} x^\varpi & \text{if } [\varrho] = [Z/\varpi] \text{ and } x \in Z \\ x_\mathbf{n} & \text{otherwise} \end{cases}$
- $(f(t_1, \ldots, t_n))^{\Gamma_\mathbf{n}[\varrho]} :\approx \begin{cases} (f(t_1, \ldots, t_n))^\varpi & \text{if } [\varrho] = [Z/\varpi] \text{ and } \\ & f(t_1, \ldots, t_n) \in Z \\ f_\mathbf{n}(t_1^{\Gamma_\mathbf{n}[\varrho]}, \ldots, t_n^{\Gamma_\mathbf{n}[\varrho]}) & \text{otherwise} \end{cases}$
- $Z^{\Gamma_\mathbf{n}[\varrho]} := \{z^{\Gamma_\mathbf{n}[\varrho]} \mid z \in Z\}$

(ii) for all $\varphi \in \mathrm{FORM}_{\mathbf{L}^*}$

- $(R(t_1, \ldots, t_n))^{\Gamma_\mathbf{n}[\varrho]} :\approx R(t_1^{\Gamma_\mathbf{n}[\varrho]}, \ldots, t_n^{\Gamma_\mathbf{n}[\varrho]})$
- $(t_1 \equiv t_2)^{\Gamma_\mathbf{n}[\varrho]} :\approx t_1^{\Gamma_\mathbf{n}[\varrho]} \equiv t_2^{\Gamma_\mathbf{n}[\varrho]}$
- $\bot^{\Gamma_\mathbf{n}[\varrho]} :\approx \bot$

- $(\varphi \to \psi)^{\Gamma_{\mathbf{n}}[\varrho]} :\approx \varphi^{\Gamma_{\mathbf{n}}[\varrho]} \to \psi^{\Gamma_{\mathbf{n}}[\varrho]}$
- $(\Box_Z \varphi)^{\Gamma_{\mathbf{n}}[\varrho]} :\approx \begin{cases} \Box_{Z^{\mathbf{n}[\text{-}]}} \varphi^{\Gamma_{\mathbf{n}}[\text{-}]} & \text{if } [\varrho] = [\text{-}] \\ \Box_{Z^{\mathbf{n}[X/\varpi]}} \varphi^{\Gamma_{\mathbf{n}}[Z \cap X/\varpi]} & \text{if } [\varrho] = [X/\varpi] \end{cases}$
- $\Phi^{\Gamma_{\mathbf{n}}[\text{-}]} := \{\varphi^{\Gamma_{\mathbf{n}}[\text{-}]} \mid \varphi \in \Phi\}$

N.B.: In a feeble effort to enhance legibility the Γ's are sometimes left out of the translation-functions in the remainder of this section. No ambiguity can possibly arise since the translation is dependent on the subscripts anyway.

The eventual aim is then to prove for all wurs $\Gamma_{\mathbf{n}}[\varrho] \in$ WUR:
$\mathcal{M}_{\mathbf{PLuM}}, \Gamma_{\mathbf{n}}[\varrho] \models \varphi \iff \varphi^{\mathbf{n}[\varrho]} \in \Gamma$.
First, however, some auxiliary results should be obtained.

Proposition 4.5 Let $\varphi((\Box_Y \psi)^{\mathbf{m}[Y/\mathbf{n}[\varrho]]})$ be a formula containing $(\Box_Y \psi)^{\mathbf{m}[Y/\mathbf{n}[\varrho]]}$ as a subformula that is not in the scope of a modal operator, then:
$\vdash \varphi((\Box_Y \psi)^{\mathbf{m}[Y/\mathbf{n}[\varrho]]}) \Longrightarrow$ for all $\mathbf{k} \in \omega : \vdash \varphi((\Box_Y \psi)^{\mathbf{k}[Y/\mathbf{n}[\varrho]]})$

Proof: By a tedious induction on the length of the derivation.[1] ⊣

The next step is the Lindenbaum lemma, showing that each consistent theory in $\mathbf{L}_\omega$ can in general be extended to a maximal consistent one. The other result shows how every consistent theory in $\mathbf{L}^*$ corresponds to an equally consistent theory in $\mathbf{L}_\omega$.

Lemma 4.6 ('Lindenbaum Lemma')

(*i*) If $\Gamma \subseteq \text{FORM}_{\mathbf{L}^*}$ is consistent, then $\Gamma^{\mathbf{n}[\text{-}]} \subseteq \text{FORM}_{\mathbf{L}_\omega}$ is also consistent.

(*ii*) For each language $L \in \mathcal{L}^*$ and $\Gamma \subseteq \text{FORM}_L$:
if Γ is consistent, i. e. $\Gamma \nvdash \bot$, there is a maximal consistent $\Gamma' \subseteq \text{FORM}_L$ extending Γ, i. e. $\Gamma' \supseteq \Gamma$ and $\Gamma' \nvdash \bot$

Proof:

(*i*) Straightforward. Merely, indexing the various lexical items with one and the same element, here $\mathbf{n}$, does not affect the deductive properties of a theory.

(*ii*) As in the propositional modal case. Let $\varphi_1, \ldots, \varphi_n, \ldots$ be an enumeration of FORM_L. Then define inductively:

- $\Gamma_0 := \Gamma$
- $\Gamma_{n+1} := \begin{cases} \Gamma_n & \text{if } \Gamma_n \cup \{\varphi_{n+1}\} \vdash \bot \\ \Gamma_n \cup \{\varphi_{n+1}\} & \text{otherwise} \end{cases}$

Finally set $\Gamma' := \bigcup_{n \in \omega} \Gamma_n$. Γ′ can then straightforwardly be shown to be both maximal and consistent. ⊣

[1] Intellectual integrity compels me to confess that I have not yet executed the allegedly tedious induction. The proposition, however, seems plausible enough. Note that no bound terms get liberated and no free terms bound by the translation. Still, in order to demonstrate the proof some induction loading seems to be in order.

We are now in a position to construct the canonical m-model $\mathcal{M}_{\mathbf{PLuM}}$ for $\mathbf{L}^*$. Consider *all* maximal consistent theories For each $\Gamma \subseteq \text{FORM}_{\mathbf{L}_\omega}$ and each boldface integer $\mathbf{n} \in \omega$, a possible world $\Gamma_{\mathbf{n}}$ will be constructed. But first define:

Definition 4.7 For each maximal consistent theory in $\Gamma \subseteq \text{FORM}_{\mathbf{L}_\omega}$ and all $\mathbf{t} \in \text{TERM}_{\mathbf{L}_\omega}$:

- $[\mathbf{t}]_\Gamma := \{\mathbf{t}' \mid \mathbf{t} \equiv \mathbf{t}' \in \Gamma\}$.

And only then define the worlds $\Gamma_{\mathbf{n}}$ of $\mathcal{M}_{\mathbf{PLuM}}$ as follows:

Definition 4.8 (The worlds of $\mathcal{M}_{\mathbf{PLuM}}$) Define for each maximal consistent theory Γ in $\mathbf{L}_\omega$ and each $\mathbf{n} \in \omega$, $\Gamma_{\mathbf{n}}$ as:

$\Gamma_{\mathbf{n}} := \langle D_{\Gamma_{\mathbf{n}}}, [\![\cdot]\!]_{\Gamma_{\mathbf{n}}}, g_{\Gamma_{\mathbf{n}}} \rangle$ with:

- $D_{\Gamma_{\mathbf{n}}} := \{[\mathbf{t}]_\Gamma \mid \mathbf{t} \in \text{TERM}_{\mathbf{L}_\omega}\}$
- $[\![\cdot]\!]_{\Gamma_{\mathbf{n}}}$ as an interpretation function of the non-logical constants as:
  - $[\![c]\!]_{\Gamma_{\mathbf{n}}} := [c_{\mathbf{n}}]_\Gamma$
  - for each n-place function symbol f: $[\![f]\!]_{\Gamma_{\mathbf{n}}} :=$ any function $\mathbf{h}$ such that: $\mathbf{h}([\mathbf{t}_1]_\Gamma, \ldots, [\mathbf{t}_n]_\Gamma) = [f_{\mathbf{n}}(\mathbf{t}_1, \ldots, \mathbf{t}_n)]_\Gamma$
  - $[\![R]\!]_{\Gamma_{\mathbf{n}}} := \{\langle [\mathbf{t}_1]_\Gamma, \ldots, [\mathbf{t}_n]_\Gamma \rangle \mid R_{\mathbf{n}}(\mathbf{t}_1, \ldots, \mathbf{t}_n) \in \Gamma\}$
- $g_{\Gamma_{\mathbf{n}}}$ such that for each $x \in \text{VAR}_{\mathbf{L}^*}$: $g_{\Gamma_{\mathbf{n}}}(x) = [x_{\mathbf{n}}]_\Gamma$

Let REST^* denote $\text{REST}^{\mathbf{L}^*}_{\mathcal{M}_{\mathbf{PLuM}}}$ for the remainder of this section. Then define the canonical m-model, $\mathcal{M}_{\mathbf{PLuM}}$ for **PLuM** with respect to $\mathbf{L}^*$ as:

Definition 4.9 The canonical m-model $\mathcal{M}_{\mathbf{PLuM}} := \langle \mathcal{W}, \mathcal{R} \rangle$ with:

- $\mathcal{W} := \{\Gamma_{\mathbf{n}} \mid \Gamma \text{ is a maximal consistent theory in } \mathbf{L}_\omega \ \& \ \mathbf{n} \in \omega\}$
- $\Gamma_{\mathbf{n}} \mathcal{R} \Gamma'_{\mathbf{m}} :\iff$
  for all $\Box_Z \varphi \in \text{FORM}_{\mathbf{L}^*}$ and all $[\varrho] \in \text{REST}^*$:
  $(\Box_Z \varphi)^{\Gamma_{\mathbf{n}}[\varrho]} \in \Gamma \Longrightarrow \varphi^{\Gamma'_{\mathbf{m}}[Z/\Gamma_{\mathbf{n}}[\varrho]]} \in \Gamma'$

$\mathcal{M}_{\mathbf{PLuM}}$ is an m-model. Demonstrating this involves a series of tedious representation independence proofs, which will be omitted here.

Fact 4.10 $\mathcal{M}_{\mathbf{PLuM}}$ is an m-model

Proof: Tedious ⊣

By contrast, the following lemma is a crucial intermediary step:

Lemma 4.11 For $\mathcal{M}_{\mathbf{PLuM}}$, for all $\Gamma_{\mathbf{n}} \in \mathcal{W}$ and all $[\varrho] \in \text{REST}^*$:

$(\Box_Z \varphi)^{\Gamma_{\mathbf{n}}[\varrho]} \in \Gamma \iff$ for all $\Gamma'_{\mathbf{m}}$ such that $\Gamma_{\mathbf{n}} \mathcal{R} \Gamma'_{\mathbf{m}}$: $(\varphi)^{\Gamma'_{\mathbf{m}}[Z/\Gamma_{\mathbf{n}}[\varrho]]} \in \Gamma'$

Proof: Consider $\mathcal{M}_{\mathbf{PLuM}} = \langle \mathcal{W}, \mathcal{R} \rangle$, and arbitrary $\Gamma_{\mathbf{n}} \in \mathcal{W}$ and $[\varrho] \in \text{REST}^*$:

$\Rightarrow$: Directly from the definition of $\mathcal{R}$.

$\Leftarrow$: Assume that $(\Box_Z \varphi)^{\mathbf{n}[\varrho]} \notin \Gamma$. Suppose further, for a *reductio ad absurdum* that there is no $\Gamma'_{\mathbf{m}} \in \mathcal{W}$ such that $\Gamma_{\mathbf{n}} \mathcal{R} \Gamma'_{\mathbf{m}}$ such that $\varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]} \notin \Gamma'$. By the Lindenbaum Lemma (lemma 4.6(*ii*), above) this means that for *all* $\mathbf{m} \in \omega$:

(1) $\{(\neg\varphi)^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}\} \cup \{\psi^{\mathbf{m}[Y/\mathbf{n}[\varrho']]} \mid (\Box_Y \psi)^{\mathbf{n}[\varrho']} \in \Gamma \ \& \ [\varrho'] \in \text{REST}^*\} \vdash \bot$

As can easily be verified, the deduction theorem holds for the axiom system of **PLuM**— just note that (Nec) can only be applied if there are no premises operative. Applying the deduction theorem and the fact that a derivation can involve a finite number of formulae only, one may subsequently obtain for all $\mathbf{m} \in \omega$ through some Boolean reasoning:

(2) $\{\psi^{\mathbf{m}[Y/\mathbf{n}[\varrho']]} \mid (\Box_Y \psi)^{\mathbf{n}[\varrho']} \in \Gamma \ \& \ [\varrho'] \in \mathrm{R{\small EST}}^*\} \vdash \varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}$

(3) $\vdash \bigwedge_{1 \leq i \leq n} \psi_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]} \rightarrow \varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}$

Because of the finiteness of the formula in (3) and all $\mathbf{n}[\varrho_i]$ being of a finite depth, there are entirely 'new' $\mathbf{m} \in \omega$ — i. e. some $\mathbf{m}$ such that no term $f_{\mathbf{m}}(\mathbf{t}_1, \ldots, \mathbf{t}_n)$ or variable $x_{\mathbf{m}}$ occurs in either any $(\Box_{Y_i} \psi_i)^{\mathbf{n}[\varrho_i]}$ $(1 \leq i \leq n)$ or $(\Box_Z \varphi)^{\mathbf{n}[\varrho]}$ — for which (3) holds. Consider any such $\mathbf{m}$. Let $F_{\mathbf{m}} = \{f_{\mathbf{m}} \mid f \in \mathrm{F{\small UNC}}_{\mathbf{L}^*}\} \cup \{x_{\mathbf{m}} \mid x \in \mathrm{V{\small AR}}_{\mathbf{L}^*}\}$. Some reflection will reveal that since $\mathbf{m}$ is entirely 'new', for each $t \in RT(\bigwedge_{1 \leq i \leq n} \psi_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]} \rightarrow \varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}) \cap \breve{F}_{\mathbf{m}}$: $t \approx f_{\mathbf{m}}(\mathbf{t}_1, \ldots, \mathbf{t}_n)$ or $t \approx x_{\mathbf{m}}$ for respectively some $f_{\mathbf{m}} \in F_{\mathbf{m}}$ or $x_{\mathbf{m}} \in F_{\mathbf{m}}$. Accordingly, (Nec) can be applied:

(4) $\vdash \Box_{\overset{\triangledown}{F}_{\mathbf{m}}} (\bigwedge_{1 \leq i \leq n} \psi_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]} \rightarrow \varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]})$

Applications of (A2), (A3) and some Boolean reasoning warrant the inference to subsequently:

(5) $\vdash \bigwedge_{1 \leq i \leq n} \Box_{\overset{\triangledown}{F}_{\mathbf{m}}} \psi_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]} \rightarrow \Box_{\overset{\triangledown}{F}_{\mathbf{m}}} \varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}$

(6) $\vdash \bigwedge_{1 \leq i \leq n} \Box_{\overset{\triangledown}{F}_{\mathbf{m}} \cap RT(\psi_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]})} \psi_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]} \rightarrow$
$\Box_{\overset{\triangledown}{F}_{\mathbf{m}} \cap RT(\varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]})} \varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}$

It can easily be verified that for each $\psi \in \mathrm{F{\small ORM}}_{\mathbf{L}^*}$ and each $Z \in \mathrm{I{\small NDEX}}_{\mathbf{L}^*}$:

$\overset{\triangledown}{F}_{\mathbf{m}} \cap RT(\psi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}) \subseteq Z^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}$

and so by (A3) and applying definition 4.4 follow:

(7) $\vdash \bigwedge_{1 \leq i \leq n} \Box_{Y_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]}} \psi_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]} \rightarrow \Box_{Z^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}} \varphi^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}$

(8) $\vdash \bigwedge_{1 \leq i \leq n} \Box_{Y_i^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]}} \psi_i^{\mathbf{m}[Y_i \cap Y_i/\mathbf{n}[\varrho_i]]} \rightarrow \Box_{Z^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}} \varphi^{\mathbf{m}[Z \cap Z/\mathbf{n}[\varrho]]}$

(9) $\vdash \bigwedge_{1 \leq i \leq n} (\Box_{Y_i} \psi)^{\mathbf{m}[Y_i/\mathbf{n}[\varrho_i]]} \rightarrow (\Box_Z \varphi)^{\mathbf{m}[Z/\mathbf{n}[\varrho]]}$

Then by proposition 4.5:

(10) $\vdash \bigwedge_{1 \leq i \leq n} (\Box_{Y_i} \psi)^{\mathbf{n}[Y_i/\mathbf{n}[\varrho_i]]} \rightarrow (\Box_Z \varphi)^{\mathbf{n}[Z/\mathbf{n}[\varrho]]}$

Investigating definition 4.4, especially the last clause, one will come to recognize that for each $\psi \in \mathrm{F{\small ORM}}_{\mathbf{L}^*}$ and each $Z \in \mathrm{I{\small NDEX}}_{\mathbf{L}^*}$: $(\Box_Z \psi)^{\mathbf{n}[Z/\mathbf{n}[\varrho]]} \approx (\Box_Z \psi)^{\mathbf{n}[\varrho]}$ and so:

(11) $\vdash \bigwedge_{1 \leq i \leq n} (\Box_{Y_i} \psi)^{\mathbf{n}[\varrho_i]} \rightarrow (\Box_Z \varphi)^{\mathbf{n}[\varrho]}$

Since for each $1 \leq i \leq n$, $(\Box_{Y_i} \psi_i)^{\mathbf{n}[\varrho_i]} \in \Gamma$, this would enable one to infer:

(12) $\Gamma \vdash (\Box_Z \varphi)^{\mathbf{n}[\varrho]}$,

which, however, is at variance with assumption that $(\Box_Z \varphi)^{\mathbf{n}[\varrho]} \notin \Gamma$. $\dashv$

The following fact ensures the objects that make up the domains of any two worlds $\Gamma_{\mathbf{n}}$ and $\Gamma'_{\mathbf{m}}$ such that $\Gamma_{\mathbf{n}} \mathcal{R} \Gamma'_{\mathbf{m}}$, are the same.

Fact 4.12 For $\mathcal{M}_{\mathbf{PLuM}}$, for all $\Gamma'_{\mathbf{m}}, \Gamma_{\mathbf{n}} \in \mathcal{W}$ and all $[\varrho] \in \text{Rest}^{\mathbf{L}^*}_{\mathcal{M}_{\mathbf{PLuM}}}$:
if $\Gamma'_{\mathbf{m}} \mathcal{R} \Gamma_{\mathbf{n}}$ then for all $\mathbf{t} \in \text{Form}_{\mathbf{L}_\omega}$: $[\mathbf{t}]_{\Gamma'} = [\mathbf{t}]_{\Gamma}$

Proof: In virtue of (A4)–(A6) and the definition of $\mathcal{R}$. By induction on $\mathbf{t}$. ⊣

As a penultimate result one can now obtain the following lemma:

Lemma 4.13 For $\mathbf{L}^*$, $\mathcal{M}_{\mathbf{PLuM}}$, and for all $\Gamma_{\mathbf{n}} \in \mathcal{W}$ and for all $t \in \text{Term}_{\mathbf{L}^*}$, all $[\varrho] \in \text{Rest}(\Gamma_{\mathbf{n}})$ and for all $\varphi \in \text{Form}_{\mathbf{L}^*}$:

(*i*) $[\![t]\!]_{\Gamma_{\mathbf{n}}[\varrho]} = [t^{\Gamma_{\mathbf{n}}[\varrho]}]_{\Gamma}$

(*ii*) $\mathcal{M}, \Gamma_{\mathbf{n}}[\varrho] \models \varphi \iff \varphi^{\mathbf{n}[\varrho]} \in \Gamma$

Proof: Consider an arbitrary $\Gamma_{\mathbf{n}} \in \mathcal{W}$ and $[\varrho] \in \text{Rest}(\Gamma_{\mathbf{n}})$.

(*i*) By induction on t. Here only the induction step will be done in full.

- $t \approx f(t_1, \ldots, t_n)$. This case is in turn by induction on $[\varrho]$.
  - $[\varrho] = [\text{-}]$. Just consider the following equations:
    $[\![f(t_1, \ldots, t_n)]\!]_{\Gamma_{\mathbf{n}}[\text{-}]} = [\![f]\!]_{\Gamma_{\mathbf{n}}}([\![t_1]\!]_{\Gamma_{\mathbf{n}}[\text{-}]}, \ldots, [\![t_n]\!]_{\Gamma_{\mathbf{n}}[\text{-}]}) =_{i.h.}$
    $[\![f]\!]_{\Gamma_{\mathbf{n}}}([t_1^{\Gamma_{\mathbf{n}}[\text{-}]}]_{\Gamma}, \ldots, [t_n^{\Gamma_{\mathbf{n}}[\text{-}]}]_{\Gamma}) =_{4.8} [(f(t_1, \ldots, t_n))^{\Gamma_{\mathbf{n}}[\text{-}]}]_{\Gamma}$
  - $[\varrho] = [Z/\Gamma'_{\mathbf{m}}[\varrho']]$. In this case there are two induction hypotheses operative; one deriving from the induction on t, *i.h.*(*i*) and the one deriving from the induction on $[\varrho]$, *i.h.*(*ii*). Two cases should be distinguished, either $f(t_1, \ldots, t_n) \in Z$ or $f(t_1, \ldots, t_n) \notin Z$. The proof of the latter case runs along much the same lines as the case $[\varrho] = [\text{-}]$ and *i.h.*(*i*) should be applied. If the former, note that since $[Z/\Gamma'_{\mathbf{m}}[\varrho']] \in \text{Rest}(\Gamma_{\mathbf{n}})$, $\Gamma'_{\mathbf{m}} \mathcal{R} \Gamma_{\mathbf{n}}$, and so consider the following equations:
    $[\![f(t_1, \ldots, t_n)]\!]_{\Gamma_{\mathbf{n}}[Z/\Gamma'_{\mathbf{m}}[\varrho']]} = [\![f(t_1, \ldots, t_n)]\!]_{\Gamma'_{\mathbf{m}}[\varrho]} =_{i.h.(ii)}$
    $[(f(t_1, \ldots, t_n))^{\Gamma'_{\mathbf{m}}[\varrho']}]_{\Gamma'} =_{4.12} [(f(t_1, \ldots, t_n))^{\Gamma'_{\mathbf{m}}[\varrho']}]_{\Gamma} =$
    $[(f(t_1, \ldots, t_n))^{\Gamma_{\mathbf{n}}[Z/\Gamma'_{\mathbf{m}}[\varrho]]}]_{\Gamma}$.

(*ii*) By induction on φ. If $\varphi \approx R(t_1, \ldots, t_n)$ or $\varphi \approx t_1 \equiv t_2$, invoke (*i*), immediately above. The case in which $\varphi \approx \bot$ is trivial and if $\varphi \approx \varphi \to \psi$ a straightforward application of the induction hypothesis will do the trick. This leaves us with the case where $\varphi \approx \Box_Z \psi$. Then just consider the following equivalences:

$\mathcal{M}_{\mathbf{PLuM}}, \Gamma_{\mathbf{n}}[\varrho] \models \Box_Z \psi$
$\iff$ For all $\Gamma'_{\mathbf{m}}$ such that $\Gamma_{\mathbf{n}} \mathcal{R} \Gamma'_{\mathbf{m}}$: $\mathcal{M}_{\mathbf{PLuM}}, \Gamma'_{\mathbf{m}}[Z/\Gamma_{\mathbf{n}}[\varrho]] \models \psi$
$\iff_{\text{i.h.}}$ For all $\Gamma'_{\mathbf{m}}$ such that $\Gamma_{\mathbf{n}} \mathcal{R} \Gamma'_{\mathbf{m}}$: $\psi^{\Gamma'_{\mathbf{m}}[Z/\Gamma_{\mathbf{n}}[\varrho]]} \in \Gamma'$
$\iff_{4.11}$ $(\Box_Z \psi)^{\Gamma_{\mathbf{n}}[\varrho]} \in \Gamma$ ⊣

Completeness now follows as a corollary.

Corollary 4.14 For all $\Gamma \cup \{\varphi\} \subseteq \text{Form}_{\mathbf{L}^*}$:
$\Gamma \models \varphi \iff \Gamma \vdash \varphi$

Proof: Consider arbitrary $\Gamma \cup \{\varphi\} \subseteq \text{Form}_{\mathbf{L}^*}$:

$\Leftarrow$: This is of course the soundness part of the proof that has already been demonstrated by theorem 4.2.

$\Rightarrow$: Adequacy of the axiom system can best be demonstrated by proving the contraposition. So suppose $\Gamma \nvdash \varphi$. Then $\Gamma \cup \{\neg\varphi\}$ is consistent and in virtue of lemma 4.6 (*i*) above $(\Gamma \cup \{\neg\varphi\})^{\Gamma_{\mathbf{n}}[\text{-}]}$ should also be consistent. As such there should be a $\Gamma_{\mathbf{n}}$ among the worlds of the canonical model $\mathcal{M}_{\mathbf{PLuM}}$ such that $(\Gamma \cup \{\neg\varphi\})^{\Gamma_{\mathbf{n}}[\text{-}]} \subseteq \Gamma$. From lemma 4.13(*ii*) follows subsequently that $\mathcal{M}_{\mathbf{PLuM}}, \Gamma_{\mathbf{n}}[\text{-}] \models \gamma$ for all $\gamma \in \Gamma$ and $\mathcal{M}_{\mathbf{PLuM}}, \Gamma_{\mathbf{n}}[\text{-}] \models \neg\varphi$. $\mathcal{M}_{\mathbf{PLuM}}$ being an m-model entails $\Gamma \not\models \varphi$, and we are done. $\dashv$

Chapter 5

Rigid Designation and Direct Reference

5.1 Rigid Designation

5.1.1 Kripke on Proper Names and Rigid Designation

One of the more general philosophical points Kripke tried to make in his *Naming and Necessity*[1] and 'Identity and Necessity'[2] is that some referring expressions need not possess a descriptive content in order to refer. Furthermore, he argued that the use of proper names in natural languages is a good case in point. On this view proper names are neither disguised definite descriptions nor disjunctions or clusters of these. Contrariwise, names would lack a reference determining sense entirely. In this respect, Kripke's semantical views evince a stark contrast with Fregean semantics. In theories of the latter kind each proper name is allotted a *Sinn* or sense, which somehow describes the object referred to. Kripke did not consider these cluster or description theories to be inconsistent, he just thought them to be false (cf. Kripke [1980], p.64). The reasons Kripke had for repudiating Fregean and Fregean-style theories of proper names will not be dealt with in this section. Kripke's view, which he derived from Mill, that:

> [A] proper name is, so to speak, *simply* a name. It *simply* refers to its bearer, and has no other linguistic function. In particular, unlike definite description, a name does not describe its bearer as possessing any special identifying properties. (Kripke [1979], pp.239-40)

will be subscribed to without much ado. I merely confine myself to giving vent to my opinion that the Millian picture of proper names and naming in natural languages is superiour to the Fregean account.

On Kripke's view the relations between definite descriptions and and proper names are not entirely severed. A definite description may be employed to fix the reference of a name, still, one should one should be careful not to confuse the

[1] Kripke [1972] and Kripke [1980]
[2] Kripke [1971]

former with the meaning of the latter. Once the reference of a proper name has been fixed, the success of its referring depends merely on the object's existence, not on its having certain properties. Circumstances may obtain in which one would conclude that a certain object does not exist; the identity conditions of an object may be satisfied or not. The descriptive content of a referring expression should be carefully distinguished from the identity conditions of the object to which it is supposed to refer. An object may not exist because no object satisfies its identity conditions. In such a case, a name meant to denote that object, fails to refer, not so much because there is no object satisfying its descriptive content — it does not have any — but merely because there is no object to refer to in the first place. A proper name may fail to refer because an object happens not to exist, but it cannot fail to refer because an object lacks certain properties. To put the point differently, if we use a name to refer to a certain object, and we find out that the object has widely different properties from the ones we initially thought it had — for the sake of argument suppose the objects lacks all properties we thought it had short from us concluding that the object does not exist — the name continues to denote the object. At some stage we may come to realize that we have been wrong in supposing that a certain object existed, and that the name has failed to refer all along. That, however, has nothing to do with the alleged meaning of the name.

Another distinction Kripke makes is between rigid and non-rigid designators. Rigid designators are referring expressions which denote the same object in each possible world.[3] Non-rigid designators are referring expressions that are not rigid designators. In this sense mathematical expressions such as '$\sqrt[3]{27}$' are usually taken to be rigid, whereas for 'the president of the United States' this is not the case as a rule. The intuitive test Kripke gives for a referring expression α to be a rigid designator or not is whether 'α could not have been α' could be true under some readings, either *de dicto* or *de re*, of the different occurrences of α. If under no reading this can be true α is a rigid designator, and a non-rigid one otherwise. For example, '$\sqrt[3]{27}$ could not have been anything else but $\sqrt[3]{27}$. Yet the actual president of the United States would not have been the president of the United States had he lost the elections. If Kripke is right, proper names are rigid designators.

At a first glance it may seem that the thesis that proper names lack a reference determining sense and the claim that proper names are rigid designators pull in opposite directions. If proper names have no descriptive content, it would seem that any name could name any object. However, if names are rigid designators, they denote the same object in all worlds, and so necessarily denote a particular object.

Although Kripke does not put it in so many words, the two claims concerning proper names may be reconciled as follows. The use of a proper name consists in singling out a particular object in order to say something about that object. The use of 'Gödel' in:

[3] To this Kripke usually adds that this should not be taken to imply that the referent of a proper name must exist in all possible worlds. Rather, rigid designators denote the same object in each possible world in which that object exists. Designators which rigidly denote an object that does exist in all possible worlds may be called *strongly* rigid. Cf. Kripke [1980], p.48, and also Kripke [1971], p.79.

(5.1) Gödel proved the incompleteness of arithmetic

is to single out Gödel and to convey that he proved the incompleteness of arithmetic. Obviously, if 'Gödel' were replaced by the name 'Tarski', (5.1) would fail to impart anything about Gödel. The mysterious thing here is how people manage to refer to the right object with a proper name assuming the latter does not have a reference determining sense. Kripke's answer is that the use of a proper name is determined by the object it denotes. Before a proper name can be used correctly, its reference should be fixed, and at that very moment its use as a rigid designator is established as well. If we talk about other possible worlds, or counterfactual situations, we should employ our language with its regular uses. In each possible world, the predicates 'cow' and 'chair' should denote cows and chairs respectively and not the other way round. If Kripke is right about proper names lacking altogether a descriptive content determining their references, their correct use is completely conditioned by the object they denote. Thus, when talking about other possible worlds proper names should be used as we use to use them, i.e. they should be deployed with the denotations they have in the actual world (cf. Kripke[1980], pp.77, 108-9, 109 n.51; Kripke[1971], p.78).

This way of putting it may suggest to the reader that Kripke's claims that proper names should always get a *de re* interpretation. It might seem that Kripke contention is that, although the distinction between *de re* and *de dicto* readings does make sense for definite descriptions, it does not for proper names. Matters are, however, subtler than just this. This interpretation does not chime with what Kripke wanted to convey, as witness his introduction to the 1980 edition of *Naming and Necessity*:

> The thesis that names are rigid in simple sentences, however, is equivalent (ignoring complications arising from the possible nonexistence of the object) to the thesis that if a modal operator governs a simple sentence containing a name, the two readings of large and small scope are equivalent. This is not the same as the doctrine that natural language has a convention that only the large scope reading is allowed. In fact, the equivalence makes sense only for a language where both readings are admissible.[4] (Kripke[1980], p.12, n.15)

Thus, Kripke maintains that the rigidity of proper names comes down to the *de dicto* and *de re* readings of modal statements containing occurrences of proper names being equivalent. He did not allege so much as *de dicto* modal statements involving proper names making no sense whatsoever.

It may be wondered what can be behind this. One way of understanding Kripke has to do with the way we speak about metaphysical possibilities, and the image of possible worlds we should entertain. The famous Morningstar paradox is a convenient example to illustrate matters. So consider:

(5.2) Hesperus is Hesperus,

(5.3) Hesperus is Phosphorus,

(5.4) Necessarily, Hesperus is Phosphorus.

[4] I take Kripke here as meaning *de dicto* with 'small scope' and *de re* with 'large scope'. For circumstantial evidence of Kripke's use in this respect compare Kripke [1979], pp.241-2. Also compare Kripke [1971], p.83 fn.10.

As has already been observed before, no object can fail to be self-identical, and that this holds true no matter how that object is referred to. So, since Hesperus is in fact identical to Phosphorus, (5.4) is true if both 'Hesperus' and 'Phosphorus' get a *de re* reading. If 'Hesperus' and 'Phosphorus' get a *de dicto* interpretation in (5.4), however, matters are slightly different. The question then concerns the circumstances in which the truth conditions of (5.3) are satisfied rather than the circumstances in which the planet Venus is identical with the planet Venus. It is important to take notice of the almost trivial observation that it is the truth conditions of the string of symbols

(5.5) $\mathrm{H}^\frown\mathrm{e}^\frown\mathrm{s}^\frown\mathrm{p}^\frown\mathrm{e}^\frown\mathrm{r}^\frown\mathrm{u}^\frown\mathrm{s}^\frown\sqcup^\frown\mathrm{i}^\frown\mathrm{s}\ ^\frown\sqcup^\frown\mathrm{P}^\frown\mathrm{h}^\frown\mathrm{o}^\frown\mathrm{s}^\frown\mathrm{p}^\frown\mathrm{h}^\frown\mathrm{o}^\frown\mathrm{r}^\frown\mathrm{u}^\frown\mathrm{s}$

as they are for the language we speak in the actual world that we are interested in, and not so much the truth conditions of (5.5) in some other language. If so, the observation that people might have spoken another language in which (5.5) is false, however true, is completely beside the point.[5] The real issue pivots on the question whether we could conceive of circumstances in which the English truth conditions of (5.3) are not satisfied.[6] If we can, in any such circumstances there should be at least two objects, one of which is denoted by 'Hesperus' and the other by 'Phosphorus'. The problem now is whether these are circumstances in which (5.3) is false or circumstances in which (5.5) has been interpreted in some other language. As has already been remarked above, there are restrictions on the use of the English lexicon. Thus, when talking about other possible worlds the English lexicon should also get its denotations in the various worlds in accordance with its ordinary use:

> But still, in describing that world, we use *English* with *our* meanings and *our* references. It is in this sense that I speak of rigid designator as having the same reference in all possible worlds. (Kripke[1980], p.77)

The question is what this entails with respect to the interpretation of proper names in other possible worlds, considering their lacking all descriptive meaning.

Since the use of a proper name is fully determined by the object it denotes, one may seek a way out of these quandaries by stipulating any language in which proper names refer to other objects is simply not the language we speak. Thus any purported counter example against the truth of (5.4), even if taken *de dicto*, is comparable to the case in which people could be said to speak a different language in which (5.3) happens to be false. Any such case, however, is not germane to the issue, as has already been observed above. This argument would undoubtedly clinch the case for the rigidity of the proper names. Still, setting definite identity criteria for natural languages is a precarious affair to say the least and the more contentious thesis that proper names are rigid designators should, in my opinion,

[5] On this point I am indebted to Robert Goijers, who called my attention to this way of presenting the argument.

[6] By employing the verb 'conceive' in this sentence I wish by no means to imply that which possible worlds there are is somehow dependent on our imaginative powers, or anything along those lines. The formulation has been chosen in order to avoid confusion with what I would have meant to convey had I written 'The issue pivots on the question whether the English truth conditions of (5.5) would have been satisfied had those circumstances be actual'. I can imagine the reader wondering, on first reading what this gibberish is all about. I hope to clarify matters when I return to the issue in the next section.

not be made dependent upon any proposals in this respect. Anyway, Kripke makes the case for the rigidity of proper names on different and rather more sophisticated grounds.

Kripke's arguments are connected with his views on the infamous transworld identity problem. This epistemological puzzle concerns the question how to identify the objects of the actual world with objects in other possible worlds. The complicating factor is that the whole idea behind the possible world metaphor is that objects, or their counterparts, may have different properties in different possible worlds. The matter concerning transworld identity or the counterpart relation should, accordingly, either be settled on grounds independent of the properties the objects have in the different possible worlds, or by imposing restrictions on the degree in which objects, or their counterparts, can be different in different possible worlds. With respect to this enigma many and widely diverging positions have been taken up and subsequently been defended by both philosophers and logicians, of whom some would undoubtedly contest the viability of my formulation of the problem.

Kripke suspects that the transworld identity problem emanates from a misguided way of looking upon possible worlds as some kind of distant planets, given to us purely qualitatively.[7] Rather, possible worlds should be conceived of as counterfactual situations, ways in which the objects of this world could have had different properties. The objects involved when evaluating a modal sentence can then be identified in the actual world by the properties they have in the actual world, before considering in what way they could have been different.

> We can refer to an object and ask what might have happened to *it*. So, we do not begin with worlds (which are supposed somehow to be real, and whose qualities, but not whose objects, are perceptible to us), and then ask about criteria of transworld identification; on the contrary, we begin with the objects, which we *have* and can identify, in the actual world. We can then ask whether certain things might have been true of the objects. (Kripke [1972], p.273)

Thus the transworld identity problem vanishes.

The spurious image of possible worlds as some kind of distant planets may also suggest that possible worlds are given to us including a certain interpretation of the language in which we wish to speak about them. And the practice of modal logic may induce one to believe that one can concoct counter examples to modal sentences by manipulating the interpretation function practically at whim. Exploiting Kripke's picture of possible worlds as counterfactual situations, however, makes clear that there is no choice as to the interpretation of the proper names. Speaking about other possible worlds from the vantage point of the actual world, names should be assigned their denotations in those worlds in conformity with their denotations in the actual world. If not, the denotations of the proper names in those other worlds would not be in accordance with their ordinary use in our world. Consequently, proper names should be conceived of as rigid designators.[8]

[7] Cf. Kripke [1971], p.81.

[8] Like Kripke I am here deliberately ignoring possibly embarrassing issues connected with possible non-existence of certain objects in other possible worlds.

It may be countered that Hesperus, i. e. the heavenly body that is Venus, could surely have been different in the sense of being called 'Mars'. The immediate rejoinder would be "being called 'Mars' by whom?". In view of the above remarks certainly not by us, stuck in our own actual world and restricted to our own linguistic practices as we are.[9] Perhaps by people, if any, inhabiting that other world, but why would that compromise *our* naming the objects in that world the way we do? The way I like to think about it, the people, using 'Mars' to denote Venus as they are wont to, may still be considered speakers of the English language if their language resembles ours to a sufficient degree in other respects. Even so, their deployment of the English language is just as irrelevant to ours when speaking about their world as had they been speaking German, or any other language for that matter. What if things had been different in the sense that there just happened to be no creatures speaking any language whatsoever? Would it be impossible to speak about that world in a *de dicto* fashion? It seems not. There certainly could be non-equivalent *de re* and *de dicto* statements involving definite descriptions. Consider, for instance:

(5.6) The highest mountain of the southern hemisphere could have been the highest peak of the world.

If both the referring expressions get a *de dicto* interpretation (5.6) seems to be true. Surely, the northern hemisphere could have been as flat as a pancake and the southern hemisphere much the same as it actually is. Moreover, these considerations are quite independent of whether or not in those counterfactual situations there would be creatures endowed with speech. Yet, if both 'the highest mountain of the southern hemisphere' and 'highest peak in the world' are taken *de re*, (5.6) imparts the possible identity of the Aconcagua and Mount Everest, which of course cannot be the case, their being entirely different mountains. When contemplating the ways things could have been different, we interpret the language we speak in those counterfactual situations, and project our interpretations of the proper names onto those worlds.

It is important to note that these reflections concern the interpretation of proper names in other possible worlds and not so much the properties the objects denoted by proper names in the actual world have in other possible worlds. Yet, the interpretation of a name in any possible world will correspond to the interpretation of that very name in the actual world, i. e. the *de dicto* interpretation of a name will coincide with its *de re* interpretation. The latter, of course, matches exactly with what Kripke maintained.

So much for the necessary truth of (5.3) both when taken *de dicto* and *de re*.[10] This, however, still leaves Kripke to account for the cognitive difference between (5.2) and (5.3). Obviously, the formal arguments of subsection 3.2.2, above, cannot be put to use, since on Kripke's account no metaphysically possible worlds exist in which proper names have a different denotation. In particular, in each possible

[9] Surely, *we* could have been different from the way we are in the sense of calling Venus 'Mars'. This, however, is not the point. What is at stake is whether given their actual uses, or given any of their counterfactual uses for that matter, proper names are rigid designators.

[10] Of course, only the *de re* variant will be an immediate consequence of the principle of indiscernibility of identicals. The *de dicto* reading will then follow from this and the assumption that proper names are rigid designators.

world both 'Hesperus' and 'Phosphorus' will refer to the planet Venus.

The solution Kripke offers is based on the conceptual difference aprioricity and aposterioricity on the one hand and necessity and contingency on the other. The former pair belongs to the domain of epistemology. Something is *a priori* true if it can be known independently of any experience, and known *a posteriori* if this is not the case. The difference between necessity and contingency, on the other hand, has to do with how things could have been different from the way they actually are. Without trying to give precise definitions, the metaphysical and the epistemic modal pair can be seen to require at least an elaborate philosophical argument before they can be matched. Kripke thought that any such argument is apt to miscarry. What is more, if his arguments are correct, not even their respective extensions coincide. So, although to recognize that 'Hesperus is Hesperus' is true merely demands a proper understanding of the notion of identity and does not require even so much as a quick glance at the heavens, it took quite some astronomical sense and talent to establish that 'Hesperus' and 'Phosphorus' are names of the same planet. The cognitive difference between 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' can thus be attributed to the fact that the former is an *a priori* and the latter only an *a posteriori* truth.

5.1.2 Some Remarks on Naming and the Epistemic Modalities *a priori* and *a posteriori*

In the previous section it was argued that proper names are rigid designators. This conclusion emanated from certain reflections on the Millian character of proper names. A consequence of the rigidity of proper names is that identity statements between proper names, of both a *de re* and a *de dicto* vintage, are either necessarily true or necessarily false. Even if the conclusion that proper names are rigid designators is granted with respect to the metaphysical modalities, the same conclusion seems less than obvious in the case of epistemic modalities. Most notably, the tenability of Kripke's solution of the Morningstar paradox depends on an asymmetry of the metaphysical and the epistemic case in this respect. However identical Hesperus and Phosphorus may be, (5.3) should still be considered an *a posteriori* truth.[11] If, however, proper names are to denote the same object in each possible world, in any of which (5.3) will accordingly be false. If, however, someone can, due to a lack of empirical information about the world, consistently believe that (5.3) is false, which possible world does he believe to be actual? No such worlds would seem to be available. One might be tempted to conclude either that any analysis of epistemic contexts in terms of possible worlds is misguided or that the Millian account of proper names should be abandoned. Kripke himself was thoroughly aware of this state of affairs, as witness:

> But although my position confirmed the Millian account of names in modal contexts, it equally appears at first blush to imply a *non-Millian* account of epistemic and belief contexts (and other contexts of propositional attitude). For I presupposed a sharp contrast between epistemic and metaphysical possibility: Before appropriate empirical discoveries were made, men might well

[11] At least in the *de dicto* case.

> have failed to know that Hesperus was Phosphorus, or even to believe it, even though they of course knew and believed that Hesperus was Hesperus.
> (Kripke [1979], p.243)

Unwilling to succumb to a non-Millian account of proper names, yet being less than satisfied with the proposed analyses of their behaviour in epistemic contexts, Kripke concluded that here we are confronted with a puzzle, on the solution of which he held no 'official doctrine'.[12]

It could be wondered what the anxiety is all about. The puzzle Kripke sees himself faced with just indicates that the cases of metaphysical and epistemic modality are asymmetrical and should be treated separately. Why should we be committed to an analysis of epistemic modalities in terms of the *metaphysically* possible worlds? What should be done, one could continue, is to distinguish between metaphysically and epistemically possible worlds. Our ignorance simply allows for more possibilities than Nature does.

This reply, however, is slightly off the mark. The point is not so much that there is a problem to account for aprioricity and aposterioricity within the formal framework meant for the analysis of the metaphysical modalities. Rather, the puzzle concerns the claim that the Millian character of proper names makes for their designating objects rigidly. Unless the arguments for the rigidity of proper names can be shown to turn on specific features of the concept metaphysical modality, one is committed to the rigidity of proper names in *any* framework that involves possible worlds. That is, if one is determined to accept the Millian view on proper names. One further remark is in place here. It is referring expressions that are rigid or non-rigid, not their usage. We are not at liberty to use expressions as rigid or as non-rigid designators. The Babylonians did not use the names 'Hesperus' and 'Phosphorus'[13] to denote different objects. They just thought they did. 'Hesperus' and 'Phosphorus' have been denoting one object, *viz.* Venus, all along. If so, it is no use contending that in metaphysical contexts we use proper names as rigid designators but that we do not in epistemic contexts.

This section will be devoted to the reconciliation of a Millian account of proper names and the aposterioricity of identity statements between proper names. The point will be that there is a difference between, on the one hand, which sentences can be said to true of certain possible but not actual circumstances and which sentences would have been true had a certain possible world been actual, on the other. The distinction may seem elusive and contrived. The reader could feel as if he were asked to distinguish between Tweedledee and Tweedledum, to use Smullyan's telling phrase. Still, I dare maintain, this distinction differentiates between the metaphysical and epistemic modal talk and opens up new vistas on the solution of Kripke's puzzle. If the argument offered here is sound, the possibility of *a posteriori* yet necessary identity statements can, moreover, be seen to sustain a Millian account of proper names rather than impairing any such venture.

When making the case for necessary *a posteriori* identity statements, I am concerned with their *de dicto* readings only. If *de re* modal statements are intelligible

[12] Cf. Kripke [1979] and Kripke [1980], p.20-1.

[13] Or whatever names they actually deployed to denote Venus. I am pretty certain that the Babylonians cannot have used the names 'Hesperus' and 'Phosphorus' at all. The names just have too Greek a ring for that.

at all, I do not think that an exception should be made if the modality involved is of an epistemic character. Consider once again Butler's dictum, every thing is what it is and no other thing. No empirical information whatsoever is needed in order to appreciate its validity. It would be utterly ridiculous, if humanly possible in the first place, to examine every object individually and check whether it is identical to itself or to any other object. Butler's Dictum is known to us *a priori*. Still, this knowledge does not merely pertain to the universal satisfaction of its truth conditions. Rather, *of* any object it is known *a priori* that it is that very object and no other. Thus it follows that if Hesperus is identical to Phosphorus it is *a priori of* Hesperus and Phosphorus that they are identical, i.e.

(5.7) Hesperus is *a priori* identical to Phosphorus

is true if both 'Hesperus' and 'Phosphorus' get a *de re* interpretation. This result may strike one as somewhat unorthodox, yet the point is fully analogous to one made before for the metaphysical case.[14] So far metaphysical and epistemic modalities behave much in the same way. Matters are different in the *de dicto* case.

In the following remarks natural languages will be considered identical if the respective uses of the lexical items apart from the proper names coincide. This, of course, is an arbitrary delineation and should be considered merely as a terminological matter. I am confident that the points to be made below could also have been made if other identity criteria for natural languages had been adopted.

In the previous section the argument that we could have spoken a language in which 'Hesperus' and 'Phosphorus' denote different objects, and thus disprove the rigidity of proper names, was weighed and found wanting. The point was that when making metaphysically modal statements we should continue using our language the way we normally do. Nevertheless, I implore the reader to reflect on this case once more. Consider a possible world, say w', in which all of us speak English except that 'Hesperus' and 'Phosphorus' designate different objects, say the latter refers to Mars whereas the former still denotes Venus. Speaking about w' from the vantage point of the actual world, w_0, in a language in which 'Hesperus' and 'Phosphorus' denote the same planet, would not impair the truth of (5.4), not even when understood *de dicto*. However, if w' had been actual we would have spoken a language in which 'Hesperus' and 'Phosphorus' are not coreferential, and neither would the truth conditions of (5.3) have been satisfied.[15] So, from the vantage point of the actual world, w' cannot be said to be different from w_0 in the sense that (5.3) is not true. The interpretation of the proper names in the actual world will be projected onto w'. Yet, when considering w' as were it the actual

[14] Again the issue of existence may be raised. Can an object be said to be self-identical if it does for some reason not exist in some possible world? Yet, this question is not specific to the epistemic case.

[15] It could be countered that the argument here depends on a particular choice as to the identity criteria of languages. The argument is bound to fail if (5.3) is taken to be ambiguous as to whether it denotes a sentence of English or a sentence of the language we would have spoken on w'. If one wishes to take this stand, and one wishes to argue for the aposterioricity of (5.3) one should also concede that it is an *a posteriori* matter which language we speak. This would be fine with me. If so, however, we would have spoken a language in which 'Hesperus' and 'Phosphorus' denote different objects, if w' had been actual. Starting from there, a similar argument as is presented here can be put forward. As such the argument does *not* depend on a particular choice as to the identity conditions of languages.

world, we no longer describe w' from an outside standpoint, such as w_0 provided in the metaphysical case, let alone we have any reason to project w_0's interpretation of the proper names onto w'.

If it can now be shown that epistemically modal statements concern what would have been the case if some other world had been actual rather than the ways things could have been different from the way they are, the epistemic puzzle concerning the possibility of *a posteriori de dicto* identity statements can be solved.

In the previous section it was argued that possible worlds should not be looked upon as some kind of distant planets, given to us purely qualitatively. I tend to concur with Kripke in this respect. *Pace* David Lewis, this picture is a spurious one indeed, apt to equip one with unsound intuitions. A better image would be to conceive of possible worlds as counterfactual situations, as ways how things could have been different. The rigidity of proper names could then be understood by means of this metaphor. So far as metaphysical modalities are concerned, this seems roughly the right picture. Yet, when dealing with epistemic modalities this way of imagining things fares about as bad as the distant planet picture. It suggests that the interpretation of the language, in particular the references of the proper names, in the other possible worlds can be determined with reference to the interpretation of the language in the actual world. With the image of possible worlds as counterfactual situations no problems ensue, since it is exactly with reference to the actual world that we conceive of the other possible worlds. Rather than talking about possible worlds as counterfactual situations, epistemic possible worlds should be construed of as worlds compatible with one's epistemic situation. Rather than ways the worlds could have been different, epistemically possible worlds are ways the world could actually be. The gist of the matter is that when dealing with epistemic modalities one lacks the outside vantage point of the actual world one has when making metaphysically modal statements. The problem is that one wishes to know which of the worlds compatible with one's information about the world *is* in fact the actual one. One thinks of possible worlds as candidates for the actual world, not as ways the world could have been but are not. Epistemic modalities concern our knowledge of how the actual world is and not so much the way the actual world could have been different.

A last remark concerns the Millian character of proper names. The difference between epistemic and metaphysical modality pertains to a difference in the way we determine the interpretation function of our language in each possible worlds. In the metaphysical case we take an outside point of view and project the interpretation of the proper names onto the other possible worlds. In the epistemic case one lacks such an outside vantage point. If so, it would seem that one, at least in principle, could speak in both metaphysical and an epistemic fashion about one particular possible world. The interpretation of proper names in that world is dependent on the point of view one takes, a metaphysical or an epistemic one. If so, a Millian account of proper names is thus in some sense vindicated. If a name had a reference determining sense, it could denote only but one object in any world, irrespective of ones metaphysical or epistemic interests.

5.1.3 Rigid Designation in PLuM

If the individual constants are to represent proper names, we saw in subsection 5.1.1 that there is good reason to take them to be rigid designators. It was also maintained that the rigidity of proper names meant so much as that the truth-value of statements is independent of whether the proper names occurring in it get a *de re* or a *de dicto* interpretation.

In **PLuM** this observation can be seen to hold. Here something slightly stronger will be proven, namely that for any rigid terms, be they atomic or not, *de re* and *de dicto* readings are equivalent. The result is not spectacularly remarkable or vexing. Actually, if you come to think about it, the claim verges on the trivial.

Definition 5.1 (Rigid Denotation) For each language $L \in \mathcal{L}$ and all $X \subseteq \textsc{Term}_L$, In any m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$:
X refers quasi-rigidly iff for all $t \in X$, for all $w, w' \in \mathcal{W}$: $[\![t]\!]_w = [\![t]\!]_{w'}$

Since **PLuM** cannot distinguish between disjoint subframes, it turns out that the equivalence of formulae in which rigid terms get a *de re* reading and those in which they get a *de dicto* reading holds in models where something slightly weaker than rigidity holds. Rigid terms need only refer quasi-rigidly to render formulae in which they get a *de dicto* reading equivalent to the ones where they get a *de re* reading.

Definition 5.2 (Quasi-Rigid Denotation) For each language $L \in \mathcal{L}$ and all $X \subseteq \textsc{Term}_L$, In any m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$:
X refers quasi-rigidly iff for all $t \in X$, for all $w, w' \in \mathcal{W}$ such that $w\mathcal{R}w'$: $[\![t]\!]_w = [\![t]\!]_{w'}$

Fact 5.3 For all $L \in \mathcal{L}$, all $X \cup \{t\} \subseteq \textsc{Term}_L$, $Z \in \textsc{Index}_L$ such that $Z \backslash X \in \textsc{Index}_L$ and all $\varphi \in \textsc{Form}_L$, and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, all $w \in \mathcal{W}$, and all $[Z/\varpi] \in \textsc{rest}(w)$:
if X refers quasi-rigidly then:
(i) $[\![t]\!]_{w[Z/\varpi]} = [\![t]\!]_{w[Z\backslash X/\varpi]}$
(ii) $\mathcal{M} \models \Box_Z \varphi \leftrightarrow \Box_{Z\backslash X} \varphi$

Proof: Consider an arbitrary m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, and an arbitrary set $X \subseteq \textsc{Term}_L$ that refers quasi-rigidly, etc.

(i) Consider an arbitrary $t \in \textsc{Term}_L$. Three cases can be distinguished:

(a) $t \in Z \cap X$
(b) $t \notin Z$
(c) $t \in Z \backslash X$.

If (a) also $t \in X$. Observe that since $[Z/\varpi] \in \textsc{rest}(w)$ there must be some $Y_1, \ldots, Y_n \subseteq \textsc{Term}_L$ and $v_0, v_1, \ldots, v_n \in \mathcal{W}$ such that $\varpi = v_0[Y_1/v_1[\ldots[Y_n/v_n]\ldots]]$. It follows, by the definition of $\textsc{rest}(w)$ (definition 3.23 on p.36) that $v_n\mathcal{R}\ldots\mathcal{R}v_0\mathcal{R}w$. This together with the rigidity of X in $\mathcal{M}$, entails that for any v_i, v_j $(0 \leq i \leq n)$: $[\![t]\!]_{v_i} = [\![t]\!]_{v_j} = [\![t]\!]_w$.

It can now be appreciated that $[\![t]\!]_{\varpi} = [\![t]\!]_w$. Consequently the following equations hold: $[\![t]\!]_{w[Z\backslash X/\varpi]} = [\![t]\!]_w = [\![t]\!]_{\varpi} = [\![t]\!]_{w[Z/\varpi]}$.

If (b) the proof is by induction on t:

- $t \in \text{ATTERM}_L$. $[\![t]\!]_{w[Z/\varpi]} = [\![t]\!]_w = [\![t]\!]_{w[Z\backslash X/\varpi]}$
- $t \approx f(t_1, \ldots, t_n)$.
  $[\![f(t_1, \ldots, t_n)]\!]_{w[Z/\varpi]} = [\![f]\!]_{w[Z/\varpi]}([\![t_1]\!]_{w[Z/\varpi]}, \ldots, [\![t_n]\!]_{w[Z/\varpi]}) =_{i.h.}$
  $[\![f]\!]_{w[Z\backslash X/\varpi]}([\![t_1]\!]_{w[Z\backslash X/\varpi]}, \ldots, [\![t_n]\!]_{w[Z\backslash X/\varpi]}) =$
  $[\![f(t_1, \ldots, t_n)]\!]_{w[Z\backslash X/\varpi]}$

If (c) : $[\![t]\!]_{w[Z/\varpi]} = [\![t]\!]_{\varpi} = [\![t]\!]_{w[Z\backslash X/\varpi]}$

(ii) By (i) we know that $[Z/\varpi] \sim^w_{RT(\varphi)} [Z\backslash X/\varpi]$ for each $\varphi \in \text{FORM}_{\mathbf{L}}$. In view of this the desired result follows as a corollary of lemma 3.34. $\dashv$

Since this result holds for any index Z, and in virtue of weak replacement (proposition 3.47, p. 47: $\models \varphi \leftrightarrow \varphi' \Longrightarrow \models \psi \leftrightarrow [\varphi/\varphi']\psi$) rigid terms $t \in X$ can be added to or subtracted from an index of a modal operator without altering the truth-value of the formula.

It may be tempting to think of $\Box_Z\varphi \leftrightarrow \Box_{Z\backslash X}\varphi$ as possibly characterizing the class of frames in which $X \subseteq \text{TERM}_L$ refers quasi-rigidly. However, since quasi-rigidity of terms is defined as a property of the interpretation function of the various worlds, it does not as such single out a class of frames. On the other hand, a model in which the not all terms $t \in X$ refer quasi-rigidly, can validate $\Box_Z\varphi \equiv \Box_{Z\backslash X}$ for all $Z \subseteq \text{TERM}_L$ by pure chance, because the interpretation of the predicate symbols can be chosen favourably. So, $\Box_Z\varphi \leftrightarrow \Box_{Z\backslash X}\varphi$ does not characterize the class of models in which X refers quasi-rigidly either. A result in this direction, however, can be obtained by introducing the concept of a Z-frame. For any model on a Z-frame the interpretation of $Z \subseteq \text{TERM}_L$ and its subterms is taken to be a part of the 'hardware'. It turns out that $\Box_Z\varphi \leftrightarrow \Box_{Z\backslash X}\varphi$ then characterizes the class of X-frames in which X refers quasi-rigidly.[16]

Definition 5.4 (Z-frames) Define for any $Z \subseteq \text{TERM}_L$ define a Z-frame as a tuple $\langle\mathcal{V}, \mathcal{R}\rangle$, where $\mathcal{R} \subseteq \mathcal{V} \times \mathcal{V}$ and $\mathcal{V}$ is a set of partially interpreted worlds $v = \langle D_v, [\![\cdot]\!]^Z_v, g^Z\rangle$. For each $v \in \mathcal{V}$, $[\![\cdot]\!]^Z_v$ and g^Z are thought of as determining the references of the atomic subterms and of the function symbols occurring in any $t \in Z$ only.

Definition 5.5 $\mathcal{M} = \langle\mathcal{W}, \mathcal{R}'\rangle$ is a model on a Z-frame $\mathcal{F} = \langle\mathcal{V}, \mathcal{R}\rangle$ iff there is an isomorphism $f : \mathcal{F} \mapsto \mathcal{M}$ such that for all $v = \langle D_v, [\![\cdot]\!]^Z_v, g^Z\rangle \in \mathcal{V}$: if $f(\langle D_v, [\![\cdot]\!]^Z_v, g^Z\rangle) = \langle D'_w, [\![\cdot]\!]_w, g\rangle$ then $D_v = D_w$ & $[\![\cdot]\!]^Z_v = [\![\cdot]\!]_w \restriction Z$ & $g^Z = g \restriction Z$. $[\![\cdot]\!]_w \restriction Z$ and $g \restriction Z$ are here respectively the interpretation function $[\![\cdot]\!]_w$ and the assignment g restricted to the interpretation of the atomic subterms of and the function symbols occurring in any $t \in Z$.

Theorem 5.6 For any $X \subseteq \text{TERM}_L$: $\Box_Z\varphi \leftrightarrow \Box_{Z\backslash X}\varphi$ characterizes the class of X-frames in which X refers quasi-rigidly, i.e. if $\mathcal{C}$ is the class of X-frames in which X refers quasi-rigidly then for all $\langle\mathcal{V}, \mathcal{R}\rangle \in \mathcal{C}$: $\langle\mathcal{V}, \mathcal{R}\rangle \models \Box_Z\varphi \iff$ for all $Z \subseteq \text{TERM}_L$: $\Box_{Z\backslash X}\varphi \iff X$ refers quasi-rigidly in $\langle\mathcal{V}, \mathcal{R}\rangle$

[16]Note that the definition of quasi-rigid denotation is for models and not for Z-frames. It is, however, almost trivial to re-define it for Z-models. So trivial in fact that it will be omitted here.

Proof: Consider an arbitrary $X \subseteq \textsc{Term}_L$.

$\Leftarrow$: Immediately by fact 5.3.

$\Rightarrow$: (sketch) Assume that $\langle \mathcal{V}, \mathcal{R} \rangle$ is not quasi-rigid with respect to X. Then there are $t \in X$ and v, v' such that $v\mathcal{R}v'$ and $[\![t]\!]_v^X \neq [\![t]\!]_{v'}^X$. Now consider a model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}' \rangle$ which is a model on $\langle \mathcal{V}, \mathcal{R} \rangle$ in virtue of an isomorphism f and for which moreover holds that for some one-place predicate symbol P, $\langle [\![t]\!]_{f(v)} \rangle \notin [\![P]\!]_{f(v')}$ as well as $\langle [\![t]\!]_{f(v')} \rangle \in [\![P]\!]_{f(v')}$. Obviously, such a model exists. Moreover, as can readily be appreciated, $\mathcal{M}, f(v) \not\models \Box_\emptyset P(t) \to \Box_{\{t\}} P(t)$ $(\approx \Box_{\{t\} \setminus X} P(t) \to \Box_{\{t\} P(t)})$. $\dashv$

5.2 Direct Reference

5.2.1 Fregean Sense vs. Direct Reference

Meaning is formally rather an elusive concept and the formal semanticist's task could be taken to put it on a sounder footing. In this respect, the notion of extension may be considered a natural point of departure, which, incidentally, may prove quite adequate for many purposes. Yet, if one also wishes to account for the cognitive, epistemic or modal differences between sentences which only differ from one another in that certain expressions are replaced by coreferential ones, extension alone fails to be entirely satisfactory. Frege, being aware of these complications, developed his theory of *Sinn* and *Bedeutung*. His ideas made quite a philosophical hit and many have followed in his wake. Here I will restrict myself to the Fregean treatment of singular denoting terms. Without claiming to be exhaustive, the following features could be taken as common to any of these Fregean theories:

(1) Each expression of a language has got an extension, its *Bedeutung*, and expresses a (Fregean) sense, *Sinn*. The sense expressed by a sentence is a proposition.

(2) The sense of an expression is a function of the senses of its constituent parts. This feature comprises the compositionality of sense. Adopting Kaplan's terminology, the semantical contribution a singular denoting expression makes to the proposition expressed by the sentence in which it occurs will be called its 'propositional component'.

(3) The sense of a singular term t is identical to the semantical features that are part of the meaning of t by virtue of which the extension of t can be determined (in each context). The sense of a singular term is said to mediate the reference of t.

Different Fregean theories may widely diverge with respect to the specific interpretation of the notion of sense. Still, propositions can conveniently be thought of as intensional semantical entities possessing a structure that roughly corresponds to that of the syntax of sentences, with properties in the place of the predicate symbols and individual concepts occupying the argument places.[17] Equally helpful it may prove to think of the way in which a Fregean sense mediates the reference of

[17]The concept of a property should here be taken in such a way numerically different properties may be coextensive and perhaps even necessarily coextensive, depending on the notion of necessity employed. On a Fregean conception of propositions the argument places should *not* be

a referential expression as comparable to a description describing objects. Thus, one could conceive of an object satisfying a Fregean sense in a similar way as of an object matching a description. It be stressed again that these remarks should not be taken as pertaining to definite definitions but rather as mere metaphors that may prove helpful mind aids. For the remainder of this subsection the concept of sense is treated as if neatly defined and well understood.

The Fregean picture of reference of singular terms can be contrasted with that of direct reference, as it has been put forward by David Kaplan (cf. especially Kaplan [1977]). On Kaplan's appreciation of direct reference, the semantical feature in virtue of which a directly referential expression denotes an object should *not* be identified with its propositional component. Accordingly, directly referential singular terms contravene (3). This does not amount to so much as saying that directly referential expressions lack a reference determining sense altogether or that there can be no semantical mechanisms operative that govern the correct use of directly referential expressions. As has been argued for by Kaplan, indexicals, both pure indexicals and demonstratives, should be classified as directly referential.[18] Nevertheless, these expressions clearly have as a part of their meaning that their reference is dependent on certain contextual features. The semantics of 'I', 'you' and 'here' require that they respectively denote the speaker, the addressee and the location of whatever context in which these expressions are used. Directly referential terms can denote different objects in different contexts of use. Moreover, these differences in reference may very well be determined by linguistic rules governing their proper use. It is merely the identification of these semantical aspects of directly referential terms and their propositional component that is denied.

The propositional component contributed by a directly referential term is supposed to be the very object it denotes in a context of use. Accordingly sentences containing directly referential terms could be thought of as behaving much in the same way as Russell understood open formulae to express propositional functions. The propositional function Russell thought to be expressed by an open formula φ, yields a proposition for each set of values for the free variables in φ. In a similar fashion, formulae containing directly referential expressions could be taken as expressing functions from the denotations of the directly referential expressions it contains to propositions. Since the denotation of a directly referential expression may vary from one context to another, it can easily be appreciated how a sentence containing a directly referential singular term may express different propositions when uttered in different contexts. Be that as it may, propositions concern the object denoted by the direct referential expressions in the actual context of use, rather than the possible denotations they could have in other contexts of use. Let α be a directly referential term. The propositional function expressed by

(5.8) α is green

yields the proposition that d has the property of being green for each denotation d of α, as it can be determined in each context. In evaluating this proposition with respect to its cognitive, epistemic or modal characteristics, certain counterfactual or epistemically possible situations may have to be investigated. In each of these

thought of as being occupied by objects. This would run counter to the Fregean conception of propositions as being composed of *senses*.

[18] Cf. Kaplan [1977] and Kaplan [1978b].

circumstances of evaluation the properties of d are relevant. The descriptive content of a directly referential expression, if any, is only relevant to determining its reference in the context of use.

The semantic image that emerges is that of directly referential expressions deriving their references from the contexts in which they are used. It is with respect to this object that the circumstances relevant to the evaluation should be investigated. This is so, even if in some of the relevant circumstances another object better complies with the conditions by means of which the reference of a directly referential expression had originally been determined in the context of utterance. The contexts in which a sentence is used and the circumstances of evaluation of a proposition play entirely different semantical roles in the evaluation of sentences containing directly referential expressions.

These reflections made Kaplan distinguish between two notions of meaning apart from extension, *viz. character* and *content*, whereas Frege only recognized one, *Sinn*. The character of an expression is that part of its meaning which determines the content of the expression in each context. The content of an expression is the semantical contribution it makes to the proposition expressed by the sentence in which the expression occurs. In particular, the content of a sentence in a certain context is the proposition it expresses itself. Note that content of non-directly referential expressions may very well be full-fledged Fregean senses. In any such case character and content coincide.

Expressions may have a *stable* character in the sense that the ways their contents are determined in each context of use are identical.[19] Note that the character of directly referential expressions may be stable in this respect. Indexicals are a case in point. 'I', e. g., will denote the speaker or agent in each context of use. Expressions may also possess a *constant* content, in which case the same contribution is made to the propositions expressed by the sentence containing them in all contexts. If an expression has got a constant content, its character can be said to be *fixed*. Clearly, the content of indexicals is not constant in this respect and their character not fixed, though stable. It is this sense that the character of an indexical can be said to be context-sensitive. The person who is the speaker may vary from one context to another, still it is the speaker itself that constitutes the propositional component of 'I' in each of these contexts. Other directly referential terms do have a constant content. The character of 'dthat(the even prime number)' (cf. subsection 5.2.7, below) is such that its content will be the integer 2 in each context. The question whether proper names should be considered equally good examples of directly referential expressions with a stable content will be tackled in subsection 5.2.9. Expressions that allow for a full Fregean treatment will possess both constant contents and stable and fixed characters. The content of such Fregean singular terms is constant because they invariably contribute the same individual concept the the propositions, although the objects that fit any such individual concept may vary from one circumstance to the other.

Recapitulating, the proponents of a semantics allowing for directly referential expressions campaigned against Frege and his allies by launching a head-on attack on (3). Ultimately the distinction between character and content was brought

[19] My use of 'stable' differs from that of Kaplan's, who seems to use it as a synonym for 'rigid' or 'fixed'

forward. This makes that all three Fregean theses are liable to revision. Granting direct reference the victory, (1)–(3) would become:

(1-1) Each expression has got both extension and character.

(1-2) In each context each expression has got a particular (propositional) content. The content of a sentence expressed in some context could be called the proposition expressed by that sentence in that context.

(2-1) The Character of the whole is a function of the character of the parts. That is, if two compound well-formed expressions differ only with respect to components which have the same Character, then the Character of the compounds is the same.[20]

(2-2) The Content of the whole is a function of the Content of the parts. That is, if two compound well-formed expressions, each set in (possibly different) contexts differ only with respect to components which *then taken in their respective contexts* have the same content, then the content of the two compounds *each taken in its own context* is the same.[20]

(3-1) In each context of use the character of a singular term determines its (propositional) content.

(3-2) In all circumstances of evaluation the content of a singular term determines its extension. This is achieved either by means of some kind of conceptual or descriptive feature, or directly. In the latter case, the character has already determined the object with respect to which the proposition expressed by the sentence has to be evaluated.[21]

5.2.2 Direct Reference and Rigid Designation

Direct reference and rigid designation have frequently been compared. Kripke argued that proper names had to be rigid designators, this being due to their lacking a reference determining sense. Using similar arguments as Kripke's proper names can be deployed showing them to be directly referential. This in itself, of course, does not render the two concepts identical. Rather, the concepts had better be kept apart carefully.

The reference of indexicals is context dependent and may vary from one context to another. In this sense sense directly referential expressions are not in general rigid designators with respect to contexts. The issue, therefore focuses on the question whether directly referential expressions are rigid designators with respect to the circumstances of evaluation and *vice versa*.

The general semantical picture of singular expressions is that of their propositional component being determined in the contexts of use through their character. In the case of a Fregean singular expression, the propositional component will be a Fregean sense. In each of the relevant circumstances of evaluation there may or

[20]These clauses are literally taken from Kaplan [1977], p.507.

[21]Semantical theories of a Fregean vintage certainly encompass rather more than just the features (1)–(3) and so does the theory of direct reference. In this section I have restricted myself to the the treatment of singular terms. Other, rather more problematic, features of the theory of direct reference, such as the treatment of natural kind terms, I have entirely disregarded. The point is, of course, that I have no solution for them ready in **PLuM**.

there may not be an object matching this Fregean sense. If there is, this object could be said to be the reference of the singular term in those circumstances of evaluation, mediated by its Fregean sense. A Fregean singular term is a rigid designator if the Fregean sense yield the same object in each of the circumstances of evaluation.

Things, as has already been emphasized, are radically different in the case of directly referential singular terms. The character of a directly referential singular term does not so much as determine a Fregean sense in each context of use, but rather an object as the propositional component. Furthermore, it is this object that is relevant to the evaluation of the proposition expressed by the sentence containing the directly referential term in the various circumstances of evaluation. The determining of this object, however, took place in the context of use and no determining the denotation of a directly referential term remains to be done in the circumstances of evaluation. So, given a context of use a directly referential term can accordingly be said, be it in a perhaps somewhat metaphorical sense, to denote its propositional component in each of the circumstances of evaluation. It is exactly in this sense of denoting that the a directly referential expression can be said to denote in each context of evaluation without the mediation of a Fregean sense. It is also in this idiosyncratic sense that a directly referential expression can be said to be a rigid designator.

There is, however, an important way in which some rigid designators are *not* directly referential. A Fregean semantics does not preclude the possibility of some singular terms being rigid designators. The propositional component contributed by singular expressions may be a Fregean sense in virtue of which they denote the same object in all possible circumstances of evaluation. 'The smallest prime number greater than nine', e. g. rigidly denotes 11 in virtue of a mediating Fregean sense, but as such is not a directly referential expression. Kripke has called such designators rigid *de facto* as to be distinguished from allegedly directly referential expressions, which would be rigid designators *de jure*[22].

Any semantics that is to provide an account of directly referential singular expressions should do some justice to this distinction between direct reference and *de facto* rigid designation. Moreover, it is a prerequisite for any such formal theory it cater for the difference between character and content.

5.2.3 Intensional Semantics à la Carnap

From a formal point of view extension is usually taken to be a rather unproblematic notion, whereas the concept of a Fregean *Sinn* is a notoriously intangible one. With respect to the identity criteria for the Fregean concept of *Sinn* several proposals have been put forward to make sense of *Sinn* in a more formal framework. Highly conspicuous are, in this respect, the attempts of Church and Carnap, who took logical equivalence as the criterion for identity of *Sinn*. This subsection will be confined to some aspects of Carnap's treatment.

Carnap endeavoured to demarcate *Sinn* by purely logical means. The criterion he employed for two expressions having the same *Sinn* was their interchangeability in intensional contexts without loss of truth-value. *Sinn* on this conception is called

[22]Cf. Kripke [1980], p.21 fn.21.

intension. The Fregean picture is completed by assigning an intension to each expression of a language. Within a formal framework comprising possible worlds or state descriptions, the intension of any expression α could be conceived of as that function that maps each world w onto the extension of α in w. If each logical possibility represented by at least one possible world, two sentences can readily be recognized as logically equivalent if and only if their intension is the same function from possible worlds to truth-values. Montague has shown, moreover how the compositional structure of intensions can be accounted for within this framework.

Notwithstanding the undeniable elegance and symmetry of this theory, the intensions allotted to directly referential terms and those of Fregean expressions are both the same kind of functions from worlds to extensions. As such, Carnap's formalization does no justice to the conceptual difference between both kinds of expression.[23] In the case of proper names this deficiency could be hushed up by taking the intension of a proper name to be a constant function. But even so, the distinction between directly referential expressions and *de facto* rigid designators gets obscured, the intensions of the latter kind of expression also being constant functions. The intension of, e. g., '11' and the one of 'the largest prime number greater than nine' will turn out to be identical. This in itself, however, cannot be held against the Carnapian project as such. If all that is deemed to be logically possible is represented by a possible world, then any two expressions that have as the same intension will be logically indiscernible. Accordingly, if one, with Carnap, wishes to demarcate *Sinn* by logical means alone, one would expect nothing else but rigid designators and directly referential expressions having got indistinguishable intensions.

Even if this is conceded, however, the Carnapian framework does do no justice to the semantical picture of direct reference as Kaplan envisaged it:

> What is characteristic of directly referential terms is that the designatum (referent) determines the propositional component rather than the propositional component, along with a circumstance, determining the designatum. [...] The propositional component need not choose its designatum from those offered by a passing circumstance; it has already been secured its designatum before the encounter with the circumstance.
>
> When we think in terms of possible world semantics this fundamental distinction becomes subliminal. This is because the style of the semantical rules obscures the distinction and makes it appear that directly referential terms differ from ordinary definite descriptions only in that the propositional component in the former case must be a *constant* function of circumstances. In actual fact, the referent, in a circumstance, of a directly referential term is simply *independent* of the circumstance and is no more a function (constant or otherwise) of circumstance, than my action is a function of your desires when I decide to do it whether you like it or not [*sic*]. The distinction that is obscured by the style of possible world semantics is dramatized by the

[23] Frege took the ideal language stance, dismissing phenomena of natural language that were obdurate enough to refuse to be straightjacketed by the Fregean doctrine as being presumably absent in an ideal formal language. Still, this rejoinder would surely not do in the case of directly referential expressions. The semantical behaviour of free variables counts as a paradigmatic case of direct reference and it would not be easy to exclude them from any ideal language the inventor of quantification theory had in mind.

structured proposition picture. That is part of the reason why I like it.
(Kaplan [1977], p.497)

Matters are even considerably worse in the case of context dependent directly referential expressions such as indexicals. The Carnapian framework has no means whatsoever to distinguish between contexts of use and circumstances of evaluation. There are only possible worlds or state descriptions to work with. Since character and content diverge in the case of indexicals one will be faced with the dilemma whether to assimilate the intension to character or to content. Putting it slightly differently, should the possible worlds be taken as contexts of utterance or as contexts of evaluation?

If the former, it may be helpful to conceive of the internal structure of the possible worlds as being augmented with some additional contextual features such as a speaker, a time, a location, turning them into what Dana Scott called 'indices'.[24] Indexical expressions such as 'I', 'here', 'now' and 'actually' could then be taken as making reference to the respective contextual features of the indices. It be noted that some restrictions on such indices are now quite in order, e.g. it should be assured that the speaker of an indexed is on the specified location at the specified time. This however would render

(5.9) I am here now

true on each of the indices. Consequently, it would seem that (5.9) expresses a necessary proposition. Although in some deep sense (5.9) should turn out to be a theorem of any logic of indexicals, this can hardly be taken as an intuitively acceptable result: surely, I could very well have had been somewhere entirely else.

If on the other hand, the context of utterance is assumed to be somehow given, and the possible worlds are to be taken as circumstances of evaluation, indexicals will behave in much the same manner as proper names do. There are then two, equally spurious courses to take. Either, one takes Dana Scott's advice to heart embellishing the worlds with speakers, hearers, locations and the lot. The context of use being assumed to be somehow given, one would expect worlds to be identical with respect to the added contextual features. In each world the same person will have to be the speaker or agent, etc. If so, one will once again find oneself at a loss accounting for the contingency of the proposition expressed by (5.9). Or one does refrain from taking indexical features into account, which, however, would amount to not giving a semantics for indexicals at all.

5.2.4 Two-Dimensional Modal Logic

In this thesis there is going to be no subsection dealing with two-dimensional modal logic in comparison with **PLuM** whatsoever.

5.2.5 Direct Reference in PLuM

The models of **PLuM** have got a two-layered structure. At the first layer are the possible worlds, at the second the worlds under a restriction, the wurs. This

[24] Cf. Scott [1970]. The indices mentioned here should, for obvious reasons, not be confused with the indices that can be attached to the modal operators in **PLuM**.

distinction between worlds and wurs can be employed to represent the difference between contexts and circumstances of evaluation and ultimately also between character and content. The aim will be to define a notion of *intension* retaining the Carnapian spirit by construing them as some kind of function from model-theoretic entities to extensions. The present analysis differs from that of Carnap's, however, in that the model-theoretic entities are taken to be a particular set of wurs rather than worlds. The worlds, as distinguished from the wurs, will play much the same role as Kaplan allotted to contexts, i. e. determining the intensions of the expressions of the language. Throughout this section no particular assumptions will be made as to the sets of directly referential terms of the languages of $\mathcal{L}$. Still, for each language L any subset of terms will do, and the treatment will be quite independent of this choice.

On Kaplan's account different contexts can yield different contents for the same expressions and as such his concept of content is relative to the context of use. In much the same vein the notion of intension will have to be rendered relative to the worlds. The wurs that are to figure as the circumstances of evaluation are relative to both the worlds and the set of directly referential expressions. Accordingly, define:

Definition 5.7 Define for each $L \in \mathcal{L}$ and for each choice of $Z \subseteq \text{TERM}_L$ as the set of directly referential expressions of L, and each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and each $w \in \mathcal{W}$, $CoE_{\mathcal{M}}(w)$ as:
$CoE_{\mathcal{M}}(w) := \{w'[Z/w] \mid w' \in \mathcal{W}\}$.
Set further $CoE_{\mathcal{M}} = \text{WUR}_{\mathcal{M}}$.
As usual, the subscript $\mathcal{M}$ will be omitted when no ambiguity arises.

The general idea is to conceive of the intension of any expression α of a language $L \in \mathcal{L}$ in a world w as that function which assigns to each wur $w'[\varrho] \in CoE$ the extension of α in $w'[Z/w]$. With respect to formulae in wurs the extension is taken to be a truth values $\mathbf{1}$ (true) or $\mathbf{0}$ (false).[25] We could adopt the following straightforward definition:

Definition 5.8 For all $L \in \mathcal{L}$, $\varphi \in \text{FORM}_L$, all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, and all $\varpi \in \text{WUR}^L_{\mathcal{M}}$:

$$[\![\varphi]\!]_{\varpi} := \begin{cases} \mathbf{1} & if\ \mathcal{M}, \varpi \models \varphi \\ \mathbf{0} & otherwise \end{cases}$$

We are now in a position to define the intension of an expression α of L in a world w, $(\![\alpha]\!)_w$, in quite a neat and general fashion as the function that assigns to each wur $w'[\varrho] \in CoE$ the extension of α in $w'[Z/w]$. Since for each $z \in Z$, $[\![z]\!]_{w'[Z/w]} = [\![z]\!]_w$, the directly referential terms derive their denotation invariably from w in all $w'[\varrho] \in CoE$. On the other hand, in each $w'[Z/w] \in CoE(w)$ the extension of each not-directly referential term $t \notin Z$ will be $[\![t]\!]_{w'}$ if t is atomic, or $[\![f]\!]_{w'}([\![t_1]\!]_{w'[Z/w]}, \ldots, [\![t_n]\!]_{w'[Z/w]})$, otherwise.

Definition 5.9 Define for each $L \in \mathcal{L}$ and for each choice of $Z \subseteq \text{TERM}_L$ as its set of directly referential expressions, each expression α of L (terms, function symbols, predicate symbols, formulae), and each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and each $w \in \mathcal{W}$ the intension of α in w, $(\![\alpha]\!)^{\mathcal{M}}_w$ as:

[25] $\mathbf{1}$ and $\mathbf{0}$ need not be added to the semantical framework as separate entities.

$(\!\![\alpha]\!\!)_w^{\mathcal{M}} :=$ that function $\mathbf{h}$ with CoE as its domain such that for each $w'[\varrho] \in CoE$: $\mathbf{h}(w'[\varrho]) = [\![\alpha]\!]_{w'[Z/w]}$

Note that the intension of an expression in a world w is also applicable to wurs $\varpi \notin CoE(w)$. The reason for this is that otherwise the same intension could not be expressed in different worlds. Moreover, it allows for the evaluation of intensions in any wur.

What Kaplan calls 'character', I will term *'pretension'*. The latter can straightforwardly be defined as a function from worlds to intensions:

Definition 5.10 For each language $L \in \mathcal{L}$, each choice of $Z \subseteq \mathrm{TERM}_L$ as its set of directly referential expressions and each expression α of L and each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, and each $w \in \mathcal{W}$ define the pretension of α in $\mathcal{M}$, $\{\!\![\alpha]\!\!\}^{\mathcal{M}}$, as:
$\{\!\![\alpha]\!\!\}^{\mathcal{M}} :=$ that function $\mathbf{h}$ with $\mathcal{W}$ as its domain such that for each $w \in \mathcal{W}$: $\mathbf{h}(w) = (\!\![\alpha]\!\!)_w$

Fixity of pretension can now be defined straightforwardly:

Definition 5.11 For any language $L \in \mathcal{L}$ the pretension of an expression α is *fixed* in an m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ iff for all $w, w' \in \mathcal{W} : \{\!\![\alpha]\!\!\}(w) = \{\!\![\alpha]\!\!\}(w')$.

5.2.6 Direct Reference in PLuM: The Picture

Having defined pretension and intension as the formal analogues of, respectively, character and content in **PLuM**, it may be wondered how well they fare. For one thing, can directly referential expressions be distinguished from *de facto* rigid designators? It should be noted that the intensions of directly referential expressions are not a different kind of entity than the intensions of terms not Z. For each context w they are both taken to be functions from the same set of wurs ($CoE(w)$) to extensions in those wurs. A first question concerns the notion of rigid reference employed. In the previous subsection definition 5.2 the notion of rigid designation was put down in model-theoretic terms. The idea was there to conceive of rigid designators as denoting the same object in all worlds of the model. What is needed here is a notion of rigid designation with respect to the $\varpi \in CoE(w)$ of the worlds ('contexts') $w \in \mathcal{W}$. Two notions, the first entailing the second, seem to be feasible:

Definition 5.12 For all $L \in \mathcal{L}$ and $t \in \mathrm{TERM}_L$ and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$

(*i*) t is a CoE-rigid designator iff for all $\varpi, \varpi' \in CoE$: $[\![t]\!]_{\varpi} = [\![t]\!]_{\varpi'}$

(*ii*) Relative to each $w \in \mathcal{W}$, t is a $CoE(w)$-rigid designator iff for all $\varpi, \varpi' \in CoE(w)$: $[\![t]\!]_{\varpi} = [\![t]\!]_{\varpi'}$

The pretension of CoE-rigid designators is fixed. To appreciate this, suppose that the pretension some term t is not fixed. Then for some m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and some $w, w' \in \mathcal{W}$ that $\{\!\![t]\!\!\}(w) \neq \{\!\![t]\!\!\}(w')$. Then $(\!\![t]\!\!)_w \neq (\!\![t]\!\!)_{w'}$, and so there should be some $w''[\varrho] \in CoE$ such that $(\!\![t]\!\!)_w(w''[\varrho]) \neq (\!\![t]\!\!)_{w'}(w''[\varrho])$. Accordingly, $[\![t]\!]_{w''[Z/w[\text{-}]]} \neq [\![t]\!]_{w''[Z/w'[\text{-}]]}$. Clearly, however, $w''[Z/w[\text{-}]], w''[Z/w'[\text{-}]] \in CoE$. Directly referential terms, i. e. $z \in Z$, on the other hand, can easily be recognized to be $CoE(w)$-rigid with respect to each $w \in \mathcal{W}$. Just observe that for any $w \in \mathcal{W}$

and any $\varpi, \varpi' \in CoE(w)$, $\varpi = w'[Z/w[\text{-}]]$ and $\varpi' = w''[Z/w[\text{-}]]$ for respectively some $w', w'' \in \mathcal{W}$, and so: $[\![z]\!]_{w'[Z/w[\text{-}]]} =_{z\in Z} [\![z]\!]_{w[\text{-}]} =_{z\in Z} [\![z]\!]_{w''[Z/w[\text{-}]]}$.

Not all CoE-rigid designators are, however, directly referential expressions. One could choose in each m-model the interpretation function in such a way that some term $t \notin Z$ is interpreted as the same object. Neither will directly referential expressions in general express an intension in a world that is distinguishable from the one expressed by a rigid designator $t \notin Z$ in the same world. It is also very well possible that the pretensions of a CoE-rigid designator $t \notin Z$ and a directly referential expression coincide. '11' and 'the smallest prime number larger than nine' will obtain the same pretension in all m-models and the same intension in each context. But how bad is this? This subsection was meant to carry out Carnap's project of demarcating the notion of sense by purely logical means in the semantical framework of **PLuM** and with reference to Kaplan's distinctions between character and content. A similar rebuttal as offered there — one would not wish to distinguish the intensions of *de facto* rigid designators and directly referential terms since they are logically very much indiscernible — seems to be in order here. Yet, it seems more important to observe that Kaplan tried to inculcate an image of the semantical workings of direct reference and to this purpose he adopted quite a metaphysical mode of expression. It is in this light that his talk of objects as being the propositional component directly referential expressions contribute to propositions should be seen. The charge against Carnap seems not to have been so much *that* the intensions of *de facto* rigid designators and directly referential expressions coincided. Kaplan's main grievance seems to be that the reason why the intension of the directly referential terms turned out to be constant functions could not be reduced to their ultimately deriving their denotations from one world, which went proxy for the context of utterance. The semantical framework of **PLuM**, however, can readily be seen to preserve the image of direct reference somewhat better. The reason why for each world w any directly referential expression $z \in Z$ turns out to be a $CoE(w)$-rigid designator is that for in the process of the interpretation of z in any $w'[Z/w] \in CoE(w)$ it is invariably and ultimately just the interpretation of z in w that of importance. For any term $t \notin Z$ there will always be some features of the interpretation-function of the other worlds $w' \in \mathcal{W}$ that will be relevant to the interpretation of t.[26]

[26]It be noted that this is still not likely to satisfy Kaplan. He was not even contented by his own formal semantics for LD in this respect. In his 'Afterthoughts' he remarks:

> The representation in possible world semantics tempts us to confuse direct reference and obstinately rigid designation. Could anyone have confused them after the clear warning of section IV? Could I have? Yes.
>
> [...] I find the confusion most evident in connection with dthat-terms, about whose syntax and semantics I seem to equivocate. [...]
>
> [...] Complete dthat-terms would be rigid, in fact *obstinately* rigid. In this case the proposition would not carry the individual itself into a possible world but rather would carry instructions to run back home and get the individual who there satisfies certain specifications. The complete dthat-term would then be a rigid description which introduces a complex 'representation' of the referent into the content; it would not be directly referential. (Kaplan [1989], pp.579-60)

I think Kaplan is here overly concerned about his beloved Russellian semantical image of direct reference. After all, it is merely an image that is instrumental to the appreciation of the underlying philosophical insight and should as such not be taken too seriously. The way I like to think about it the Russellian picture of objects being the propositional component and directly referential terms carrying instructions to run back home to obtain their denotation there serve

5.2.7 Modal Propositions

De re and *de dicto* Propositions

So far the set of directly referential terms has been assumed to be fixed for each language. No particular assumptions whatsoever have been made as to the terms in this set, although individual constants and variables may be thought to qualify as viable candidates. Having the set of terms divided into a directly referential and not-directly referential part, however, seems less than obvious. It seems that at least some singular expressions can be used both as directly referential and as descriptive phrases. Consider for instance:

(5.10) The President of the United States resides in the White House.

There seems to be some ambiguity as to the proposition expressed by (5.10). The referential phrase 'the president of the United States' could be taken in its descriptive sense, in which case (5.10) expresses a proposition which pertains to a (true) fact of American polity. Alternatively, (5.10) could be taken in such a way that the propositional component contributed by 'the President of the United States' is the incumbent president himself. On the latter understanding, 'the President of the United States' is directly referential. Assuming a rather tendentious mode of expression, it could be said that in the latter case (5.10) expresses a *de re* and in the former a *de dicto* proposition.

In a similar vein, Kaplan observed that any singular descriptive phrase could be converted into a directly referential one. For this purpose he designed 'dthat' as a kind of surrogate demonstrative with descriptive phrases going proxy for demonstrations. For any singular description ∂, 'Dthat[∂]' is supposed to be the demonstrative whose demonstration is completed by ∂ and that refers directly to the object that satisfies ∂ in the context of utterance. 'Dthat' could alternatively be introduced as a functor operating on terms. The *de re* reading of (5.10) could then also be represented as:

(5.11) Dthat[the President of the United States] resides in the White House.

To the languages of $\mathcal{L}$ 'dthat' could be added as a one-place function symbol, such that, for each $Z \in \textsc{Index}_L$, dthat$(t) \in Z$ and such that, for each world w, $[\![\text{dthat}]\!]_w$ is the identity function mapping each object onto itself.

Instead of introducing functional constants, like 'dthat', to the formal framework the whole idea of a fixed set of directly referential terms could be abandoned and the notion of intension could be made relative to a variable set of terms. The intension of an expression α in a world w relative to a set of terms, $(\![\alpha]\!)_{Z,w}$, is then the intension of α in w with the accompanying instruction that the terms in Z are to be taken directly referential. Philosophical considerations may induce one to assume that certain terms should be an element of any index and that other terms should always be thus excluded. Nevertheless, it is not unnatural to identify $\textsc{Index}_L$ with the set of sets relative to which the intensions of expressions can be determined in a language L. Accordingly, intensions could be redefined as follows:

just as well as *images* of direct reference. What I deem to be essential is that it is features of the context of use only that are relevant to the semantical interpretation of directly referential terms.

Definition 5.13 For each $L \in \mathcal{L}$ and each $Z \in \textsc{Index}_L$, each expression α of L (terms, function symbols, predicate symbols, formulae), and each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and each $w \in \mathcal{W}$, define the intension of α relative to Z, $(\![\alpha]\!)^{\mathcal{M}}_{Z,w}$, as:

$(\![\alpha]\!)^{\mathcal{M}}_{Z,w} :=$ that function $\mathbf{h}$ with CoE as its domain and such that for all $w'[\varrho] \in CoE$: $\mathbf{h}(w'[\varrho]) = [\![\alpha]\!]_{w'[Z/w[\text{-}]]}$

The notion of the pretension of an expression α and the conditions under which this can be said to be fixed can likewise be made relative to the indices:

Definition 5.14 For each language $L \in \mathcal{L}$, each $Z \in \textsc{Index}_L$ each expression α of L and each m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, and each $w \in \mathcal{W}$ define the pretension of α in $\mathcal{M}$, $\{\![\alpha]\!\}^{\mathcal{M}}_{Z}$, as:

$\{\![\alpha]\!\}^{\mathcal{M}}_{Z} :=$ that function $\mathbf{h}$ with $\mathcal{W}$ as its domain such that for each $w \in \mathcal{W}$: $\mathbf{h}(w) = (\![\alpha]\!)_{Z,w}$

Definition 5.15 For any language $L \in \mathcal{L}$ and all $Z \in \textsc{Index}_L$ the pretension of an expression α relative to Z is *fixed* in an m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ iff for all $w, w' \in \mathcal{W} : \{\![\alpha]\!\}_Z\,(w) = \{\![\alpha]\!\}_Z\,(w')$.

Two Grades of Modal Involvement

Propositions have frequently been considered the bearers of truth and necessity. On this view, truth and the modalities are most naturally to be conceived of as meta-linguistic properties. Accordingly, 'Nec' and 'True' could be introduced as two-place meta-linguistic relations between worlds and propositions expressed in contexts (worlds), with respect to an index. A feasible definition of the truth predicate can easily be obtained using the notions that are already at our disposal in **PLuM**:

Definition 5.16 For all $L \in \mathcal{L}$, all $\varphi \in \textsc{Form}_L$ and all $Z \in \textsc{Index}_L$, and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and all $w \in \mathcal{W}$ and $\varpi \in CoE$:

$\text{True}((\![\varphi]\!)_{Z,w}, \varpi) :\Longleftrightarrow (\![\varphi]\!)_{Z,w}\,(\varpi) = \mathbf{1}$

It should be noted 'True' is not just a stylistic variant of '$\models$'. The predicate 'True' allows one to evaluate the truth of a proposition expressed by a formula in some world w in some entirely different circumstance than $w[\text{-}]$. On the other hand, '$\models$' allows one to evaluate formulae more generally in wurs. There will not in general be a world $w \in \mathcal{W}$ such that for all $\varpi \in \textsc{wur}$:

$\text{True}((\![\varphi]\!)_{Z,w}, \varpi)$ iff $\mathcal{M}, \varpi \models \varphi$.

Counterexamples can again be constructed along similar lines as in the proof for fact 3.20 on page 33 above. The relations $\models$ and True do coincide, however, in a certain number of cases:

Proposition 5.17 For all languages $L \in \mathcal{L}$, all $\varphi \in \textsc{Form}_L$, all $Z \in \textsc{Index}_L$ and in all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, all $w \in \mathcal{W}$:

(*i*) $\text{True}((\![\varphi]\!)_{Z,w}, w[\text{-}]) \iff \mathcal{M}, w[\text{-}] \models \varphi$

(*ii*) $\text{True}((\![\varphi]\!)_{Z,w}, w'[\varrho]) \iff \mathcal{M}, w'[Z/w[\text{-}]] \models \varphi$

Proof: The proof is straightforward. For (ii) observe the following equivalences:

$\mathrm{True}(\llparenthesis\varphi\rrparenthesis_{Z,w}, w'[\varrho])$
$\iff \quad \llparenthesis\varphi\rrparenthesis_{Z,w}\,(w'[\varrho]) = \mathbf{1}$
$\iff_{\text{def. 5.13}} \quad [\![\varphi]\!]_{w'[Z/w[\text{-}]]} = \mathbf{1}$
$\iff \quad \mathcal{M}, w'[Z/w[\text{-}]] \models \varphi$

For (i), consider the special case of (ii) in which $w'[\varrho] = w[\text{-}]$. Observing that $[Z/w[\text{-}]] \sim^{w}_{RT(\varphi)} [\text{-}]$ and applying lemma 3.34 concludes the proof. ⊣

The intensional semantics of **PLuM**, as they have been developed so far, do not account for propositions expressed in wurs. It could be argued that this is how it should be and that doing otherwise would involve the misconception of obliterating the conceptual distinction between contexts of use and circumstances of evaluation. The wurs, however, are essential to the interpretation of modal operators in **PLuM**. Moreover, wurs cannot in general be reduced to worlds. Yet, if modality is maintained to attach to propositions, it is not clear how proper sense can be made formulae which comprise nestings of modal operators. If $\Box_Z$ is to represent 'Nec' in the object language, it is only reasonable to assume that, if the proposition expressed by φ in w with respect to Z can be ascribed necessary truth, i. e. if $\mathrm{Nec}(\llparenthesis\varphi\rrparenthesis_{Z,w}, w[\text{-}])$, then also $\mathrm{True}(\llparenthesis\Box_Z\varphi\rrparenthesis_{X,w}, w[\text{-}])$ or at least $\mathcal{M}, w[\text{-}] \models \Box_Z\varphi$. This can easily be achieved if the set of circumstances relevant to the evaluation of the propositions expressed by a formula in some context is related to the accessibility relation between the worlds. So define subsequently:

Definition 5.18 For all languages $L \in \mathcal{L}$, all $\varphi \in \textsc{Form}_L$, all $Z \in \textsc{Index}_L$ and in all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$, all $w \in \mathcal{W}$:

$$CoE_{\mathcal{R}}(w[\varrho]) := \{w'[\varrho'] \in CoE \mid w\mathcal{R}w'\}$$

Definition 5.19 For all $L \in \mathcal{L}$, all $Z \in \textsc{Index}_L$, all $\varphi \in \textsc{Form}_L$ and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$, all $w \in \mathcal{W}$ and all $w'[\varrho] \in \textsc{wur}_{\mathcal{M}}$:

$$\mathrm{Nec}(\llparenthesis\varphi\rrparenthesis_{Z,w}, w'[\varrho]) :\iff \text{ for all } \varpi \in CoE_{\mathcal{R}}(w') : \llparenthesis\varphi\rrparenthesis_{Z,w}\,(\varpi) = \mathbf{1}$$

By simply writing things out in full, one may obtain:

$$\mathrm{Nec}(\llparenthesis\varphi\rrparenthesis_{Z,w}, w[\text{-}]) \iff \text{ for all } w' \text{ such that } w\mathcal{R}w' : \mathcal{M}, w'[Z/w[\text{-}]] \models \varphi.$$

Employing the semantical clauses of **PLuM** gives:

$$\mathrm{Nec}(\llparenthesis\varphi\rrparenthesis_{Z,w}, w[\text{-}]) \iff \mathcal{M}, w[\text{-}] \models \Box_Z\varphi.$$

Finally, invoking proposition 5.17, from the above the following can be inferred:

$$\text{for each } X \in \textsc{Index}_L : \ \mathrm{Nec}(\llparenthesis\varphi\rrparenthesis_{Z,w}, w[\text{-}]) \iff \mathrm{True}(\llparenthesis\Box_Z\varphi\rrparenthesis_{X,w}, w[\text{-}])$$

Complications set in, however, if φ is a modal formula. If $\Box$ should be reducible to 'Nec', how is '$\mathrm{Nec}(\llparenthesis\Box_Z\varphi\rrparenthesis_{X,w'}, w[\text{-}])$' to be understood? Evaluating $\llparenthesis\Box_Z\varphi\rrparenthesis_{X,w'}$ in $w[\text{-}]$ with respect to its modal properties, would involve checking whether application of $\llparenthesis\Box_Z\varphi\rrparenthesis_{X,w'}$ to each $\varpi \in CoE_{\mathcal{R}}(w)$ yields $\mathbf{1}$, i. e. whether in each $w'[\varrho] \in CoE_{\mathcal{R}}(w) : \mathcal{M}, \varpi \models \Box_Z\varphi$. If $\Box$ genuinely represents 'Nec' in the object-language, then once more some proposition expressed by φ should be evaluated in the circumstances possible relative to $w'[\varrho]$, i. e. relative to $CoE_{\mathcal{R}}(w'[\varrho])$. φ can express different propositions in different worlds. The problem is that it is not clear with which of the propositions that can be expressed by φ this proposition should be identified. $\llparenthesis\varphi\rrparenthesis_{Z,w'}$ will not in general do since it would do no justice to the alleged direct reference of any terms $t \in RT(\varphi) \cap Z \cap X$, which should obtain their denotation from w and not from w'. Neither would $\llparenthesis\varphi\rrparenthesis_{Z\cap X,w}$ since then any terms $t \in RT(\varphi) \cap Z$ but $t \notin X$ would get interpreted in the various $\varpi \in CoE(w')$ instead of in w'. Again, the impossibility to reduce wurs to worlds when other languages then those in $\mathcal{L}^{**}$ are concerned each initiative towards ingenuity in this respect will be rendered void. Remarkably enough, it is Quine's second grade of modal involvement, as it is concerned with the representation of modalities as sentential operators that can be nested, that causes semantical unrest here.

It would thus seem that matters can be resolved only by (conservatively) extending the notion of intension in such a way that also sense can be made of the intension of expressions in wurs. The point is not that this cannot be achieved, as the following definitions and results indicate:

Definition 5.20 For all $L \in \mathcal{L}$, $Z \in \textsc{Index}_L$, all expressions α of L and all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, $\varpi \in \textsc{wur}_{\mathcal{M}}$:

$(\![\alpha]\!)_{Z,\varpi} :=$ that function $\mathbf{h}$ with CoE as its domain such that for all $w'[\varrho] \in CoE$: $\mathbf{h}(w'[\varrho]) = [\![\alpha]\!]_{w'[Z/\varpi]}$

Definition 5.21 For all $L \in \mathcal{L}$, all $\varphi \in \textsc{Form}_L$ and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and all $\varpi, \varpi' \in CoE$:

$\text{True}((\![\varphi]\!)_{Z,\varpi'}, \varpi) :\Longleftrightarrow ((\![\varphi]\!)_{Z,\varpi'})(\varpi) = \mathbf{1}$

Definition 5.22 For all $L \in \mathcal{L}$, all $Z \in \textsc{Index}_L$, all $\varphi \in \textsc{Form}_L$ and for all m-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, all $w \in \mathcal{W}$ and all $w'[\varrho] \in \textsc{wur}_{\mathcal{M}}$:

$\text{Nec}((\![\varphi]\!)_{Z,w}, w'[\varrho]) :\Longleftrightarrow$ for all $\varpi \in CoE_{\mathcal{R}}(w')$: $((\![\varphi]\!)_{Z,w})(\varpi) = \mathbf{1}$

As can easily be verified, the following facts can be derived from these definitions:

Fact 5.23 (i) $\text{True}((\![\varphi]\!)_{\varpi}, \varpi) \iff \mathcal{M}, \varpi \models \varphi$

(ii) $\text{Nec}((\![\varphi]\!)_{Z,\varpi}, \varpi) \iff \text{True}((\![\Box_Z \varphi]\!)_{X,\varpi}, \varpi) \iff \mathcal{M}, \varpi \models \Box_Z \varphi$

Proof: The proofs turn out to be analogous to the ones for the special cases in which $\varpi = w[\text{-}]$, which are given throughout this section. ⊣

The question is much rather whether these definitions obscure the difference between contexts of utterance and circumstances of evaluation and, if so, whether therewith the different semantical roles Kaplan allotted to circumstances and context get confused as well. As regards the first part of the question I think it can hardly be denied that if the notion intension is extended as far as in definition 5.20 the wurs double as contexts of use and circumstances of evaluation. Still, as the definition stands, the separate tasks of determining the intension and the extension of an expression in the wurs are neatly kept apart. For one, the intension of an expression could be determined in some wur, but this does not impede the possibility of its being evaluated with respect to an entirely different one. As to the evaluation of propositions, it seems one just has to abandon the idea that in evaluating formulae containing nested modal operators, one can first establish the proposition it expresses in the context of utterance for once and for all, and then check whether the proposition thus expressed is true in all circumstances. Rather, in checking formulae containing nested modal operators for their truth one is in a position in which one has to evaluate different propositions in different wurs. If one really wishes to cling to the idea that a sentence expressed in a certain context should express one meaning which is relevant to its evaluation in all different possible, actual and counterfactual circumstances, intension just does not seem to meet the demands in this respect. This is of course not saying that no such notion could be defined.

$\models$, however, does obscure the between the different semantical roles context and circumstance. This is very clear from the equivalence between $\mathrm{True}(\llparenthesis\llbracket\varphi\rrbracket\rrparenthesis_{X,\varpi}, \varpi)$ and $\mathcal{M}, \varpi \models \varphi$: $\models$ cannot represent truth of propositions expressed by formulae in some wur in some other wur. But then, of course, $\models$ stems from the Tarskian tradition that takes truth to be a property of sentences rather than of propositions.

A more positive conclusion can be obtained by focusing on the equivalence of $\mathrm{Nec}(\llparenthesis\llbracket\varphi\rrbracket\rrparenthesis_{Z,\varpi}, \varpi)$ and $\mathcal{M}, \varpi \models \Box_Z \varphi$. It shows how intimately the notions of direct reference and *de re* modality are interrelated.

5.2.8 Prospects for a Logic of Indexicals

To the lexicon of the languages of $\mathcal{L}$ a set of indexical singular terms could be added.[27] In order give a proper semantical account for these additional singular terms the worlds can be granted some additional structure. To the worlds some more parameters could be added, such as a speaker, an addressee, a location, a point in time etc.

> For more general situations one must not think of the $i \in I$ as anything as simple instants of time or even possible worlds. In general we will have:
>
> $i = (w, t, p, a, \ldots)$
>
> where the index i has many *coordinates*: for example, w is a *world*, t is a *time*, $p = (x, y, z)$ is a (three-dimensional) *position* in the world, a is an *agent*, etc. All these coordinates can be varied, possibly independently, and thus affect the truth-values of statements which have indirect references to these coordinates. (Scott [1970])[28]

In order to get things right it is imperative to align the interpretation of the various parameters, e. g. a should be located at p at the time t etc. The indexical referential expressions can then as their interpretation the obvious indexical features, e. g. $\llbracket I \rrbracket_i = a$ for each $i \in \mathcal{W}$. Suppose further that there are some predicates in the language such that lay down these restrictions, e. g. a two-place predicate 'Located' such that at each index, or as I prefer to call them, context, i, 'Located(I,here)' will be validated.

Thus increasing the structure of the possible worlds as Dana Scott advised proves quite inadequate in a Carnap-style possible worlds semantics. Most notably 'Located(I,here)' would turn out to express a necessary proposition, which it should not do. If **PLuM** is augmented with these additional indexical features no such problems ensue. That is, provided that indexical singular terms are taken to be

[27] This subsection, by no means, has as its objective to develop a fully worked out logic for indexicals. It is just meant to show how some problems that rear their ugly head in standard possible world semantics, can be avoided in **PLuM**. Accordingly, certain technical issues and difficulties will almost entirely be ignored.

[28] Also quoted in Kaplan [1978b], p.67, and in Kaplan [1977], p.508.

directly referential expressions when they occur in a sentence. 'Located(I,here)' will not in general express a necessary proposition, although being validated in each i[-] and thus being a truth of the logic of indexicals ($\models$ Located(I,here)). Counterexamples disproving that 'Located(I,here)' holds at each $i'[Z/i] \in CoE(i)$ can easily be constructed.

Example 5.24 For the sake of this example the worlds, now called indices, of m-models are taken to be quintuples $\langle D_i, [\![\cdot]\!]_i, g, \mathrm{A}_i, \mathrm{P}_i \rangle$ where $\mathrm{A}_i \in D_i$ and $\mathrm{P}_i \in D_i$ are construed as the agent and the location of the context of use i represents. Consider any such m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ with $\mathcal{W} = \{i_1, i_2\}$ and i_1, i_2 such that:

(i) $D_{i_1} = D_{i_2} = \{d_1, d_2, l_1, l_2, l_3\}$

(ii) $\mathrm{A}_{i_1} = d_1$; $\mathrm{A}_{i_2} = d_2$

(iii) $\mathrm{P}_{i_1} = l_1$; $\mathrm{P}_{i_2} = l_2$

(iv) $\mathcal{R} = \mathcal{W} \times \mathcal{W}$

Taking the restrictions on the interpretation of the indexicals 'I' and 'here' into consideration, let it further be the case that:

(v) $[\![\mathrm{I}]\!]_{i_1} = d_1$; $[\![\mathrm{I}]\!]_{i_2} = d_2$

(vi) $[\![\mathrm{here}]\!]_{i_1} = l_1$; $[\![\mathrm{here}]\!]_{i_2} = l_2$

(vii) $[\![\mathrm{Located}]\!]_{i_1} = \{\langle d_1, l_1\rangle, \langle d_2, l_2\rangle\}$; $[\![\mathrm{Located}]\!]_{i_2} = \{\langle d_1, l_3\rangle, \langle d_2, l_2\rangle\}$

It is easy to verify that both $\mathcal{M}, i_1[\text{-}] \models$ Located(I,here) and $\mathcal{M}, i_2[\text{-}] \models$ Located(I,here), and, accordingly, $\mathcal{M} \models$ Located(I,here). Still, it is *not* the case that Nec($(\![\mathrm{Located(I,here)}]\!)_{\{\mathrm{I,here}\},i_1}, i_1[\text{-}]$). Since $\mathcal{M}, i_2[\{\mathrm{I,here}\}/i_1[\text{-}]] \not\models$ Located(I,here), also $[\![\mathrm{Located(I,here)}]\!]_{i_2[\{\mathrm{I,here}\}/i_1[\text{-}]]} = \mathbf{o}$. With $i_2[\{\mathrm{I,here}\}/i_1[\text{-}]] \in CoE_{\mathcal{R}}(i_1)$, this makes that $(\![\mathrm{Located(I,here)}]\!)_{\{\mathrm{I,here}\},i_1}(i_2[\{\mathrm{I,here}\}/i_1[\text{-}]]) = \mathbf{o}$, i. e. $\mathcal{M}, i_2[\{\mathrm{I,here}\}/i_1] \not\models$ Located(I,here).

Using the same example, the intension of the indexicals 'I' and 'here' can be seen to be constant functions. $(\![\mathrm{I}]\!)_{\{\mathrm{I,here}\},i_j}$ $(j \in \{1,2\})$ is that function that maps each $i'[\varrho] \in CoE$ onto $[\![\mathrm{I}]\!]_{i'[\{\mathrm{I,here}\}/i_j]} = [\![\mathrm{I}]\!]_{i_j}$. Yet, $[\![\mathrm{I}]\!]_{i_1} \neq [\![\mathrm{I}]\!]_{i_2}$ and consequently $(\![\mathrm{I}]\!)_{\{\mathrm{I,here}\},i_1} \neq (\![\mathrm{I}]\!)_{\{\mathrm{I,here}\},i_2}$ This proves the pretension of the indexicals not to be fixed in general.

5.2.9 Proper Names

Taking cue from the soundness of the Millian view on proper names, Kripke argued that they are rigid designators with respect to counterfactual though metaphysically possible situations (cf. section 5.1.1, above). Accordingly, they would have to be classified as expressions with a fixed character (cf. section 5.2.6). It goes almost without saying that Kripke thought of proper names in way that would make Kaplan classify them as directly referential expressions, their reference allegedly not being mediated by a Fregean *Sinn*. Lucid though Kripke's arguments may be, there still seems to be room for controversy over the fixity of the character of proper names. As I represented Kripke's position, proper names would have to behave as rigid designators, since the way we speak about metaphysical possible worlds, we (should) project our interpretations of the proper names onto the counterfactual situations. Still, how can this contention be distinguished from the view that when investigating other metaphysical possibilities, we impose the restriction on the interpretation of proper names that they denote the same object there as in the actual world? On the latter image, metaphysical modality would seem to be more naturally analysed as modal operators indexed with a set of terms comprising at least all proper names. Whatever be the case, Kripke's arguments, at least the way I presented them in section 5.1.1, seem to fit this picture just as well as rigid designation, which he preferred. If so, however, proper names would never be bound and, consequently, always obtain a *de re* reading. This, in turn, contravenes Kripke's characterization of the rigidity of proper names as their *de re* and *de dicto* reading being generally equivalent.

Kaplan chimed in with Kripke on the issue of the fixed character proper names would have:

> Although it is true that two utterances of 'Aristotle' in different contexts may have different contents, I am inclined to attribute this difference to the fact that distinct homonymous words were uttered rather than a context sensitivity in the character of a single word 'Aristotle'. Unlike indexicals like 'I', proper names really are ambiguous. The causal theory of reference tells us, in terms of contextual features (including the speaker's intentions) which word is being used in a given utterance. Each such word is directly referential (thus it has a fixed content), and it also has a fixed character. Therefore, in the case of proper name words, all three kinds of meaning — referent, content and character — collapse. In this proper names are unique. They have the direct reference of indexicals, but they are not context-sensitive. Proper name words are like indexicals [in] that you can carry away from their original context without affecting their content. Because of the collapse of character, content, and referent, it is not unnatural to say of proper names that they have no meaning other than their referent. (Kaplan [1977], p.562)

I tend to demur. The way Kaplan presents his position suggests that there is an issue as to whether the same string of symbols naming different individuals in different contexts is a case of mere homonymy or just an alternative use of the same name. But surely there is none. How could one ever decide the matter; what arguments would one accept as settling the issue? Whether proper names possess a fixed character should not be made dependent on the outcome, if any ever, of so spurious an issue.

But even if this point were conceded, Kaplan's argument seems to be shaky. Suppose, for the sake of argument, that one name can in fact be used to denote the different object in different contexts. Kaplan suggests that, if so, there should be features in the respective contexts explaining the divergent uses of the name. But surely, saying that one proper name may be used to refer to different objects in different contexts, is something different from saying that the reference of the proper name is dependent on certain features of the context of use. Kaplan might therefore be said to be jumping to conclusions. Utter lack of meaning would serve just as well as an explanation. If there are no linguistic rules that determine that one should use a name in a certain context to denote some object or another, why should one use that name only to denote one particular object? (Note that this is exactly where I started my exposition of Kripke's views on proper names, above.)

One could also hope to settle the case for the alleged fixity of the character of proper names by pleading allegiance to Kripke's semantical view that the proper use of a name is established by an initial baptism and is passed in a causal chain from one link to another. The situation in which the baptism took place could be assumed to be (possibly) different from any of the contexts of use and thus be just as relevant to any of them. If, however, one persists in taking the stance that the same name can denote different individuals, one could allege that several fixings of the reference of the same name can occur, only one of which is relevant to the use of that proper name in each context. If so, the situation in which the reference of a proper name is fixed could be taken to be part of the context of use to which that particular baptism is relevant.[29]

Against the fixity of the character of proper names could be brought in that it would render coreferential names synonymous, and that this should not be, since proper names could not be said to alike in meaning because there is no meaning to be alike. The problem with this is that, by the same token, the respective meanings of two proper names cannot be said to differ either.

It appears that Kripke's and Kaplan's insights in the semantics of proper names merely pertain to their being directly referential terms and their consequent possession of a constant content in each context. Whether Millian proper names is served best by a fixed or by a feeble character, however, remains a pretty moot question. Still, if we wish our (formal) semantics also to be able to account for epistemic modal phenomena, such as the cognitive difference between 'Hesperus = Phosphorus' and 'Hesperus = Hesperus', feeble character may come in handy.

---

[29] One could wish to account for the circumstances in which proper names obtain their denotations in one's semantics. These circumstances should then be differentiated from the contexts of use and the circumstances of evaluation that have already been distinguished. To avoid confusion call these circumstances in which initial baptism take place *inaugurations*. Since it is invariably just one inauguration that is relevant to the determining of the content of a proper name in a context, for each proper name an inauguration could be associated with each context. In **PLuM** contexts could then be represented as a special kind of wurs looking something like $w[Z/w_1[Z\backslash\{c_1\}/w_2[Z\backslash\{c_1,c_2\}/w_3[\ldots]]]]$ where w is the original context, $Z=\{c_1,c_2,c_3,\ldots\}$ the set of lexical items representing the proper names and their respective $w_1,w_2,w_3,\ldots$ their inaugurations. An additional advantage of any such approach is that it enables one to account for proper names not denoting an object in the domain of reference or quantification. Certainly, whenever the reference of a proper name is fixed the object denoted by the name should exist in the domain of the inauguration. However, the domain of the inauguration and the domain of the context may be chosen to be different. A drawback, on the other hand, is that only a finite number of inaugurations can be taken into account.)

If so, proper names could be considered incomplete expressions which must be completed by an act of baptism in each inauguration. In this respect proper names could then be compared with demonstratives, which need to be completed by a demonstration in each context. I should be remarked in this respect that there seem certain restrictions to the use of proper names in inaugurations. *Pace* Johnny Cash, 'Sue' is a girl's name. All men called 'Kofi' in Ghana are born on a Friday. The name 'Benjamin' could even be contended to behave in inaugurations in a similar fashion as indexicals do in contexts. In each inauguration 'Benjamin' should denote the youngest son of a family. It certainly would be strange to christen the first-born of a pair of twins 'Benjamin' and the last-born 'Methusalem'. (It could be wondered whether this hint of oddity indicates a linguistic mistake having been made. Still this is not really the point.) Note that 'Benjamin' need not denote the last born son of a family in each context. This explains why the name 'Benjamin' could also be in use in the times when family planning had still to be invented. One should be wary not to confuse these aspects of the meaning of proper names with their character or their content.

5.2.10 A Last Remark on Epistemic Modalities

Kaplan denied that there could be operators on character in English (cf. Kaplan [1977], p.510). Alleged examples of such operators, such as 'In each context it is true that ...', he called 'monsters'. They would at best tacitly introduce mechanisms that make that the sentence in the scope of the operator is merely mentioned and not so much used. This is certainly bad news for anyone who wishes to account for the Morningstar paradox by arguing against the fixity of the character of proper name.

Within the framework of **PLuM** the proposition expressed by a formula φ in a certain context, or even wur, has been made relative to a set terms that are to be taken as directly referential in φ. If this set is taken to be empty, even if the formula contains expressions that are normally considered to be directly referential, effect can be attained that are similar to those operators on character are meant to produce. For instance, if the individual constants h and p are not taken to be CoE-rigid, there may be a wur $w[\varrho] \in CoE$ such that $\mathrm{Nec}(\llbracket h \equiv p \rrbracket_{X,w}, w[\varrho])$, if $h, p \in X$, but *not* $\mathrm{Nec}(\llbracket h \equiv p \rrbracket_{\emptyset,w}, w[\varrho])$ (check this). This might suggest that epistemic modalities could be treated as operating on intentions (content) in **PLuM**. Even apart from Kaplan's objections against monsters, with respect to the aprioricity of Butler's Dictum and the aposteriority of 'Hesperus' and 'Phosporus' both denoting Venus, this is something one would like (I leave this remark in rather a cryptic state).

Even so, I do not think that any such treatment of epistemic modalities will be quite satisfactory. The doubt a Babylonian may have as regards the identity of Hesperus and Phosphorus involves an uncertainty with respect to the context he is in. When using the names 'Hesperus' and 'Phosphorus', these words refer to a certain planet, to Venus as the case is. And so, if he were to utter some sentence conveying that Hesperus is not identical to Phosphorus, he does express a complete proposition, *viz.* the absurd proposition, which yields falsity in each circumstance of evaluation. His epistemic problem concerns his ignorance as to which proposition he has uttered in the context he is in and not so much the truth and falsity of the proposition actually expressed in other counterfactual situations. For all he is worth, he could be in any of at least two situations σ_1 and σ_2. His problem is then that, say, if uttered in σ_1 the proposition expressed by 'Hesperus is Phosphorus' would be true in (the circumstance associated with) σ_1, but the proposition expressed by 'Hesperus is Phosphorus' would be false in (the circumstance associated with) σ_2 if uttered in σ_2. Accordingly, epistemic modality may concern the the truth-value of *several* propositions as they are expressed by one sentence in the various possible contexts of utterance, in the respective circumstances associated with them. The propositions that are relevant in this respect may be determined by means of the character of the sentences the epistemic anxiety is about. Still, this is something completely different from saying that epistemic modalities operate on character. Or, is it not? How monstrous these epistemic contexts are!

Chapter 6

A Tool Box of Referential Contraptions

6.1 Quantification Theory in PLuM

As can hardly remain unnoticed, the principles governing the use of $\forall$ bear some important resemblances to those of modal operators in most common modal logics. Necessitation of $\forall$ (if φ is derivable, so is $\forall x\varphi$) and its distributivity over material implication ($\forall x(\varphi \to \psi) \to (\forall x\varphi \to \forall x\psi)$) spring to the eye. Also at a semantical level of analysis the clauses for the universal quantifier and $\Box$ exhibit conspicuous similarities. This section concerns the modal analysis of classical quantification theory (from now on **CQL=**) within **PLuM**.

One of the underlying ideas of this thesis is the contention that modal operators can bind terms in their scope in a way analogous to that of quantifiers binding variables. So far, putting it this way has merely been part of the heuristics of this paper. As things turn out, however, the binding of variables by quantifiers could be conceived of as a special case of modal operators binding terms. As such, the quantifiers can be given a modal interpretation in **PLuM**. Some restrictions on the models and the indices of the languages, however, are apposite.

Note that for any m-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and each $\langle D, [\![\cdot]\!], g \rangle$, $\langle D, [\![\cdot]\!] \rangle$ is a full model of **CQL=**, and g is a suitable assignment. Conversely, to each **CQL=** model and assignment pair corresponds a possible world that can be part of an m-model. In **CQL=** formulae are evaluated with respect to a model structure and an assignment of objects to the variables. The Tarskian semantical clause for the universal quantifier, with $g[x/d]$ being the function that assigns $g(y)$ to each $y \not\approx x$ and d to x:

$$\mathcal{A}, g \models \forall x\varphi :\Longleftrightarrow \text{ for all } d \in D_{\mathcal{A}}\text{: } \mathcal{A}, g[x/d] \models \varphi,$$

indicates that various assignments are relevant to the evaluation of formulae involving quantifiers. If the model-assignment pairs are now conceived of as possible worlds of an m-model, the Tarskian clause for the universal quantifier can be taken to concern m-models with $\langle \mathcal{A}, g \rangle$ as worlds. It is then imperative that for each $h \in \mathrm{Ass}_{\mathcal{A}}$, $\langle \mathcal{A}, h \rangle$ be a world of the m-model as well. Warranting this, each model structure $\mathcal{A}$ of **CQL=** corresponds to an m-model of **PLuM**, $\mathcal{M}_{\mathcal{A}}$, as defined below:

Definition 6.1 Define for each **CQL=**-model $\mathcal{A}$ the q-model $\mathcal{M}_\mathcal{A} = \langle \mathcal{W}, \mathcal{R} \rangle$ with:

- $\mathcal{W} = \{\langle \mathcal{A}, g \rangle \mid g \in \text{ASS}_\mathcal{A}\}$
- $\mathcal{R} = \mathcal{W} \times \mathcal{W}$[1]

The scheme is to find a suitable language $L \in \mathcal{L}$ such that for each $Z \in \text{INDEX}_L$, $\Box_Z$ behaves in any m-model $\mathcal{M}_\mathcal{A}$ as the universal quantifier of **CQL=** in the model structure $\mathcal{A}$. The non-logical lexicon of each predicate logical language corresponds could also be the lexicon of a language $L \in \mathcal{L}$. The language of $\mathcal{L}$ we are looking for can thus be presumed to share the same non-logical lexicon. Much the same applies to the Boolean operators. Accordingly, the issue focuses the indices of L.

If the predicate logical language does not contain function symbols setting $\text{INDEX}_L = \{\text{TERM}_L \backslash \{x\} \mid x \in \text{VAR}_L\}$ would produce the desired result. For each model structure $\mathcal{A}$, any two worlds of the m-model $\mathcal{M}_\mathcal{A}$ differ from one another only with respect to their respective variable assignment. The indices $\text{TERM}_L \backslash \{x\}$ assure that the wurs that need to be investigated in order to evaluate a formula $\Box_{\text{TERM}_L \backslash \{x\}} \varphi$ in a world $\langle \mathcal{A}, g \rangle$ in $\mathcal{M}_\mathcal{A}$, i. e. *all* $\langle \mathcal{A}, h \rangle [\text{TERM}_L \backslash \{x\} / \langle \mathcal{A}, g \rangle]$ since $\mathcal{R}$ is universal, differ only with respect to the interpretation of x. Moreover, any such wur can readily be recognized to be elementary equivalent to $\langle \mathcal{A}, g[x/h(x)] \rangle$. This is exactly what we want because for each object $d \in D_\mathcal{A}$ there is a world $\langle \mathcal{A}, h \rangle$ in $\mathcal{M}_A$ such that $h(x) = d$.

Matters are slightly more complicated when the language contains function symbols. Any modal operator $\Box_{\text{TERM}_L \backslash \{x\}}$ only binds those occurrences of x that do not appear as a subterm of an occurrence of a functional term. In order to avoid this, all terms containing x as a subterm should also be excluded from the index. On the other hand, one should be careful that in doing so variables distinct from x do not get bound as well. The index $\{\bar{\breve{x}}\}$, as defined in accordance with definition 3.57 on page 51, below, happens to effectuate exactly this:

Definition 6.2 For all languages $L \in \mathcal{L}$:

(*i*) $\{\breve{x}\} := \{t \in \text{TERM}_L \mid x \in RT(t)\}$

(*ii*) $\{\bar{\breve{x}}\} := \text{TERM}_L \backslash \{\breve{x}\}$

Lq, as defined immediately below, can then seen to behave on the models $\mathcal{M}_\mathcal{A}$ as a predicate logical language on the respective and corresponding **CQL=** models $\mathcal{A}$.

Definition 6.3 (**Lq**) Define for each predicate logical language with the non-logical lexicon L, **Lq** as that language in $\mathcal{L}$ such that L also constitutes the entire non-logical lexicon of **Lq** and:

[1] Requiring the set of worlds to be closed under suitable assignments here, facilitates a modal analysis of classical first-order quantification theory, which we know to be undecidable. Johan van Benthem (Benthem, J. F. A. K. van [1996], chapter 9) has shown how standard poly-modal models, with assignments as states, provide a modal foundation for predicate logic. Moreover, it turned out that 'lighter' modal variants of this poly-modal system could be obtained by weakening the condition that *all* assignments should be accounted for in the models. Some values of some variables could be dependent upon one another. It seems that a similar project could be executed with **PLuM** providing the modal foundations for first-order predicate logic. This, however, would carry too far beyond the scope of this thesis.

$$\textsc{Index}_{\mathbf{Lq}} := \{\{\breve{\overline{x}}\} \mid x \in \textsc{Var}_{\mathbf{Lq}}\}$$

As can easily be verified $\mathbf{Lq} \in \mathcal{L}^*$.

First-order predicate logic can now be mimicked in **PLuM**. Let $L_{\mathbf{CQL=}}$ be a language of classical quantified predicate logic. Then define a translation function $*$ from $L_{\mathbf{CQL=}}$ to $\mathbf{Lq}$ that just replaces all universal quantifiers $\forall x$ with a modal operator $\Box_{\{\breve{\overline{x}}\}}$.

Definition 6.4

- $t^* :\approx t$
- $R(t_1, \ldots, t_n)^* :\approx R(t_1^*, \ldots, t_n^*)$
- $(t_1 \equiv t_2)^* :\approx t_1 \equiv t_2$
- $\bot^* :\approx \bot$
- $(\varphi \to \psi)^* :\approx \varphi^* \to \psi^*$
- $(\forall x \varphi)^* :\approx \Box_{\{\breve{\overline{x}}\}} \varphi^*$

Note that $*$ is a one-to-one map of $L_{\mathbf{CQL=}}$ onto $\mathbf{Lq}$. Consequently, the converse of $*$, here denoted by $\circ$, can be used as a translation function from $\mathbf{Lq}$ to $L_{\mathbf{CQL=}}$:

Definition 6.5

- $t^\circ :\approx t$
- $R(t_1, \ldots, t_n)^\circ :\approx R(t_1^\circ, \ldots, t_n^\circ)$
- $(t_1 \equiv t_2)^\circ :\approx t_1 \equiv t_2$
- $\bot^\circ :\approx \bot$
- $(\varphi \to \psi)^\circ :\approx \varphi^\circ \to \psi^\circ$
- $(\Box_{\{\breve{\overline{x}}\}} \varphi)^\circ :\approx \forall x \varphi^\circ$

Fact 6.6

- For all $\varphi \in \mathbf{Lq}$: $\varphi \approx (\varphi^\circ)^*$
- For all $\varphi \in \mathbf{L_{CQL=}}$: $\varphi \approx (\varphi^*)^\circ$

Proof: Easy. $\dashv$

As an intermediary step the following lemma is in order.

Lemma 6.7 For $\mathbf{Lq}$, all $x \in \textsc{Var}_{\mathbf{Lq}}$, $t \in \textsc{Term}_{\mathbf{Lq}}$, $\varphi \in \textsc{Form}_{\mathbf{Lq}}$ and all q-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, for all $\langle \mathcal{A}, g \rangle \in \mathcal{W}$ and all $h \in \textsc{Ass}_{\mathcal{A}}$:

(i) $[\![t]\!]_{\langle \mathcal{A}, h \rangle [\{\breve{\overline{x}}\} / \langle \mathcal{A}, g \rangle [\text{-}]]} = [\![t]\!]_{\langle \mathcal{A}, g[x/h(x)] \rangle [\text{-}]}$

(ii) $\mathcal{M}, \langle \mathcal{A}, h \rangle [\{\breve{\overline{x}}\} / \langle \mathcal{A}, g \rangle [\text{-}]] \models \varphi \iff \mathcal{M}, \langle \mathcal{A}, g[x/h(x)] \rangle [\text{-}] \models \varphi$

Proof: : Consider arbitrary $x \in \textsc{Var}_{\mathbf{Lq}}$, $t \in \textsc{Term}_{\mathbf{Lq}}$ and $\varphi \in \textsc{Form}_{\mathbf{Lq}}$, and an equally arbitrary q-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $\langle \mathcal{A}, g \rangle$ and $h \in \textsc{Ass}_{\mathcal{A}}$.

(i) Before commencing the real proof, observe that:

($*$) for all $t \in \{\breve{\overline{x}}\}$: $[\![t]\!]_{\langle \mathcal{A}, g \rangle [\text{-}]} = [\![t]\!]_{\langle \mathcal{A}, g[x/h(x)] \rangle [\text{-}]}$

This can be demonstrated by a pretty straightforward inductive argument. Just observe that in the case $t \approx y \in \text{VAR}$ by the assumption that $y \in \{\overline{\breve{x}}\}$, $y \not\approx x$ and so clearly $g(y) = g[x/h(x)](y)$. Furthermore, take notice of the fact that if $f(t_1, \ldots, t_n) \in \{\overline{\breve{x}}\}$, also for each t_i $(1 \leq i \leq n)$, $t_i \in \{\overline{\breve{x}}\}$. Consequently, the induction hypothesis is applicable.

The remainder of the proof is again by induction on t.

- $t \in \text{CONS}_L$. Observe that $c \in \{\overline{\breve{x}}\}$, and so the following equations hold:
  $[\![c]\!]_{\langle \mathcal{A},h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\text{-}]]} = [\![c]\!]_{\langle \mathcal{A},g\rangle[\text{-}]} =_{(*)} [\![c]\!]_{\langle \mathcal{A},g[x/h(x)]\rangle[\text{-}]}$.
- $t \approx y \in \text{VAR}_L$. Two cases should be distinguished; either $y \approx x$ or $y \not\approx x$. If the latter, the case is analogous to the one of t being an individual constant. If the former, however, $y \notin \{\overline{\breve{x}}\}$, and then:
  $[\![x]\!]_{\langle \mathcal{A},h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\text{-}]]} = [\![x]\!]_{\langle \mathcal{A},h\rangle[\text{-}]} = h(x) = g[x/h(x)](x) =$
  $[\![x]\!]_{\langle \mathcal{A},g[x/h(x)]\rangle[\text{-}]}$.
- $t \approx f(t_1, \ldots, t_n)$. Again two cases should be distinguished. Either $f(t_1, \ldots, t_n) \in \{\overline{\breve{x}}\}$ or $f(t_1, \ldots, t_n) \notin \{\overline{\breve{x}}\}$. The latter case can be clinched by invoking the induction hypothesis. If the former the following equations hold:
  $[\![f(t_1, \ldots, t_n)]\!]_{\langle \mathcal{A},h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\text{-}]]} = [\![f(t_1, \ldots, t_n)]\!]_{\langle \mathcal{A},g\rangle[\text{-}]} =_{(*)}$
  $[\![f(t_1, \ldots, t_n)]\!]_{\langle \mathcal{A},g[x/h(x)]\rangle[\text{-}]}$.

(ii) (sketch) Analogously to lemma 3.34 it could be proven that if two wurs coincide on the interpretation of the non-logical lexicon and they have access to the same sets of worlds in the respective model, they are elementary equivalent, i. e. they validate the same formulae. With (i) it can readily be appreciated that this is the case for any pair $\langle \mathcal{A}, g[x/h(x)]\rangle$ and $\langle \mathcal{A}, h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A}, g\rangle]$. ⊣

As a final result the following can be obtained:

Theorem 6.8 For all **CQL=**-models $\mathcal{A}$ and $g \in \text{ASS}_{\mathcal{A}}$:

(i) For all $\varphi \in \mathbf{L}_{\mathbf{CQL=}}$: $\mathcal{A}, g \models_{\mathbf{CQL=}} \varphi \iff \mathcal{M}_{\mathcal{A}}, \langle \mathcal{A}, g\rangle[\text{-}] \models_{\mathbf{PLuM}} \varphi^*$

(ii) For all $\varphi \in \mathbf{Lq}$: $\mathcal{M}_{\mathcal{A}}, \langle \mathcal{A}, g\rangle[\text{-}] \models_{\mathbf{PLuM}} \varphi \iff \mathcal{A}, g \models_{\mathbf{CQL=}} \varphi^\circ$

Proof: Consider arbitrary **CQL=**-models $\mathcal{A}$ and $g \in \text{ASS}_{\mathcal{A}}$:

(i) First it should be observed that it is trivial to show that for any $t \in \text{TERM}_{\mathbf{Lq}}$: $[\![t^*]\!]_{\langle \mathcal{A},g\rangle} = [\![t]\!]^{\mathcal{M}_{\mathcal{A}}}_{\langle \mathcal{A},g\rangle}$. The remainder of the proof is by induction on φ, of which only the case $\varphi \approx \Box_Z \psi$ is worth treating in full:

⇒: Suppose $\mathcal{M}_{\mathcal{A}}, \langle \mathcal{A}, g\rangle \not\models (\forall x\varphi)^* \approx \Box_{\{\overline{\breve{x}}\}}\varphi^*$. Then there should be some $h \in \text{ASS}_{D_{\langle \mathcal{A},g\rangle}}$ such that $\langle \mathcal{A}, g\rangle \mathcal{R} \langle \mathcal{A}, h\rangle$ and $\mathcal{M}_{\mathcal{A}}, \langle \mathcal{A}, h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A}, g\rangle[\text{-}]] \not\models \varphi^*$. By proposition 6.7 also: $\mathcal{M}_{\mathcal{A}}, \langle \mathcal{A}, g[x/h(x)]\rangle[\text{-}] \not\models \varphi^*$. Now, the induction hypothesis can be applied in order to obtain: $\mathcal{A}, g[x/h(x)] \not\models \varphi$. Since, $h(x) \in D_{\mathcal{A}}$, $\mathcal{A}, g \not\models \forall x\varphi$

$\Leftarrow$: This direction of the proof is also by contraposition, so assume $\mathcal{A}, g \not\models \forall x\varphi$. Then there should be some $d \in D_{\mathcal{A}}$ such that $\mathcal{A}, g[x/d] \not\models \varphi$. Since we are dealing with q-models, in which no suitable assignment is left unaccounted for, there must be some $h \in \text{Ass}_{\mathcal{A}}$ such that $h(x) = d$. Subsequently applying the induction hypothesis and proposition 6.7 will get us: $\mathcal{M}_{\mathcal{A}}, \langle\mathcal{A}, h\rangle[\{\overline{\overline{x}}\}/\langle\mathcal{A}, g\rangle[\text{-}]] \not\models \varphi^*$. Since clearly $\langle\mathcal{A}, g\rangle\mathcal{R}\langle\mathcal{A}, h\rangle$, the remainder is straightforward.

(ii) Almost immediately from fact 6.6 and (i), above. $\dashv$

6.2 Adding Quantifiers to PLuM

The previous section would warrant a bi-modal analysis of quantified modal logic, with one of the modalities understudying quantification. Although the spirit of such an analysis will be retained in this section, the presentation will diverge somewhat. The quantifiers will be added with their usual syntax to the languages of $\mathcal{L}$, but they will obtain the modal interpretation of the previous section. So, define the set of languages for quantified modal logic $\mathcal{K}$ as follows:

Definition 6.9 The lexicon L of each language $L \in \mathcal{K}$ is the lexicon of a language $L' \in \mathcal{L}$ with $\forall$ added as a logical constant. The syntax of each language $L \in \mathcal{L}q$ is also like in definition 3.3, on page 23 above, in every respect but for the following additional clause:

(vi) If $\varphi \in \text{Form}_L$ then $\forall x\varphi \in \text{Form}_L$

The treatment of quantified modal logic in this section will be confined to languages $\mathcal{K}^*$ with the indices of the modal operators closed under taking subterms. $\exists x\varphi$ is the usual abbreviation of $\neg\forall x\neg\varphi$. The semantical clause for $\forall x$ can then be defined as:

Definition 6.10 For all $L \in \mathcal{K}$, and all $\varphi \in \text{Form}_L$ and all m-models $\mathcal{M} = \langle\mathcal{W}, \mathcal{R}\rangle$, and all $\langle\mathcal{A}, g\rangle[\varrho] \in \text{wur}_{\mathcal{M}}$:

- $\mathcal{M}, \langle\mathcal{A}, g\rangle[\varrho] \models \forall x\varphi :\Longleftrightarrow$ for all $h \in \text{Ass}_{\mathcal{A}} : \mathcal{M}, \langle\mathcal{A}, h\rangle[\{\overline{\overline{x}}\}/\langle\mathcal{A}, g\rangle[\varrho]] \models \varphi$

In this definition mention of the accessibility relation for the modal operator $\forall$ is suppressed. It is assumed to hold between any two worlds that differ only with respect to their variable assignments and between such worlds only. The desired results are attained if the analysis is properly restricted to a special subset of all m-models. For any m-model $\mathcal{M} = \langle\mathcal{W}, \mathcal{R}\rangle$ in this subset it should be the case that, if $\langle\mathcal{A}, g\rangle \in \mathcal{W}$, then for all $h \in \text{Ass}_{\mathcal{A}}$, $\langle\mathcal{A}, h\rangle \in \mathcal{W}$. Moreover, to prevent merely alphabetically distinct formulae from being non-equivalent, any two worlds that only differ from one another with respect to the values they assign to the variables should access, as well as being accessed from, the same possible worlds. Accordingly, we obtain the notion of an mq-model with respect to which quantified **PLuM** can be investigated:

Definition 6.11 (mq-models) An mq-model is an m-model $\langle\mathcal{W}, \mathcal{R}\rangle$ such that:

(i) For all $h, g \in \text{Ass}_D$: if $\langle D, [\![\cdot]\!], g\rangle \in \mathcal{W}$ then $\langle D, [\![\cdot]\!], h\rangle \in \mathcal{W}$, and (*ii*) For all $w = \langle D_w, [\![\cdot]\!]_w, g\rangle, w' = \langle D_{w'}, [\![\cdot]\!]_{w'}, h\rangle$: if $w\mathcal{R}w'$, $g' \in \textsc{Ass}_{D_w}$ and $h' \in \textsc{Ass}_{D_{w'}}$ then $\langle D_w, [\![\cdot]\!]_w, g'\rangle\mathcal{R}\langle D_{w'}, [\![\cdot]\!]_{w'}, h'\rangle$

The restriction to mq-models compromises necessitation in quantified **PLuM**. Although

$$\models \varphi \Longrightarrow \models \Box_{\emptyset}\varphi$$

and:

$$\models \varphi \Longrightarrow \models \Box_Z\varphi$$

for all indices Z of the respective language, still hold good in quantified **PLuM**, proposition 3.58 on page 52 above, does not. This is due to the fact that in the proof of proposition 3.58 an construction on m-models is used that does not in general transform an mq-model into another mq-model. Consequently, the inference from $\models \exists x\ x \equiv a$ to $\models \Box_{\{a\}}\exists x\ x \equiv a$ is blocked. On the account of quantification as presented here, entities do not exist necessarily. The domains of the various worlds of mq-models need not coincide nor grow when accessible from one another. The following example illustrates these points:

Example 6.12 Consider an m-model $\mathcal{M} = \langle\mathcal{W}, \mathcal{R}\rangle$ with $\mathcal{W}$ containing only two worlds $w_1 = \langle\mathcal{A}, g\rangle$ and $w_2 = \langle\mathcal{B}, h\rangle$. Let $\mathcal{R} = \{\langle w_1, w_2\rangle\}$. Let further the domains of w_1 and w_2 be respectively the singletons $\{d_1\}$ and $\{d_2\}$ with $d_1 \neq d_2$. Consequently, $[\![a]\!]_{w_1} = d_1$ and $[\![a]\!]_{w_2} = d_2$. Moreover, g and h must assign respectively d_1 and d_2 to all variables of the language and are as such the only suitable assignments for $\mathcal{A}$ and $\mathcal{B}$ respectively. So $\mathcal{M}$ is an mq-model as well. $\mathcal{M}, w_1[\text{-}] \not\models \Box_{\{a,x\}}\exists x\ x \equiv a$, since
$[\![x]\!]_{\langle\mathcal{B},h\rangle[\{\bar{\breve{x}}\}/\langle\mathcal{B},h\rangle[\{a,x\}/w_1[\text{-}]]]} =_{x\notin\{\bar{\breve{x}}\}} h(x) = d_2 \neq d_1 = [\![a]\!]_{w_1} =_{a\in\{\bar{\breve{x}}\}}$
$[\![a]\!]_{\langle\mathcal{B},h\rangle[\{\bar{\breve{x}}\}/\langle\mathcal{B},h\rangle[\{a,x\}/w_1[\text{-}]]]}$
and w_2 being accessible from w_1 and h being the only assignment for $\mathcal{B}$.

Example 6.12 also provides a counterexample against the Barcan formula ($\forall x\Box_Z\varphi \to \Box_Z\forall x\varphi$). Just let $Z = \{a, x\}$ and $\varphi \approx x \equiv a$.[2] The converse of the Barcan formula $\Box_Z\forall x\varphi \to \forall x\Box_Z\varphi$ can also easily be obtained by letting $[\![P]\!]_{w_2} = \{d_2\}$. Then, $\mathcal{M}, w_1 \not\models \Box_{\{x\}}\forall xPx \to \forall x\Box_{\{x\}}Px$, as can easily be verified. These results are not very remarkable since the domains of the worlds of mq-models are required to shrink nor to grow.

A quantifier $\forall x$ (or $\exists x$) can bind a variable inside the scope of a modal operator $\Box_Z$ from without only if $x \in Z$. If not, x will be bound by the modal operator, rendering it unreachable for the quantifier. To the evaluation of the formula $\forall x\Box_Z\varphi$, with $x \notin Z$, in a wur $\langle\mathcal{A}, g\rangle[\varrho]$, the values of $[\![t]\!]_{w'[Z/\langle\mathcal{A},g'\rangle[\{\bar{\breve{x}}\}/\langle\mathcal{A}g\rangle[\varrho]]]}$ for each $t \in RT(\varphi)$ with respect to all $g' \in \textsc{Ass}_{\mathcal{A}}$ and all w' such that $\langle\mathcal{A}, g\rangle\mathcal{R}w'$ will be relevant. Having restricted the analysis to languages $\mathcal{K}^*$, and so $Z \subseteq \{\bar{\breve{x}}\}$, it turns out that for all w' such that $\langle\mathcal{A}, g\rangle\mathcal{R}w'$ and for all $g' \in \textsc{Ass}_{\mathcal{A}}$ $[\![t]\!]_{w'[Z/\langle\mathcal{A},g'\rangle[\{\bar{\breve{x}}\}/\langle\mathcal{A},g\rangle[\varrho]]]} = [\![t]\!]_{w'[Z/\langle\mathcal{A},g\rangle[\varrho]]}$.

As a pendant to lemma 6.7, above, the following result can be obtained for mq-models and languages $L \in \mathcal{K}^*$

Lemma 6.13 For all $L \in \mathcal{K}^*$,$x \in \textsc{Var}_L$, $\varphi \in \textsc{Form}_L$, and for all mq-models $\mathcal{M} = \langle\mathcal{W}, \mathcal{R}\rangle$, $\langle\mathcal{A}, g\rangle \in \mathcal{W}$, and $h \in \textsc{Ass}_{\mathcal{A}}$:

[2] Counterexamples to the case where Z is opaque, e. g. $Z = \emptyset$ are even easier to find.

(i) $[\![t]\!]_{\langle\mathcal{A},h\rangle[\{\breve{\bar{x}}\}/\langle\mathcal{A},g\rangle[\varrho]]} = [\![t]\!]_{\langle\mathcal{A},g[x/h(x)]\rangle[\varrho]}$

(ii) $\mathcal{M},\langle\mathcal{A},h\rangle[\{\breve{\bar{x}}\}/\langle\mathcal{A},g\rangle[\varrho]] \models \varphi \iff \mathcal{M},\langle\mathcal{A},g[x/h(x)]\rangle[\varrho] \models \varphi$

Proof: Largely analogous to, though somewhat more complicated than, the proof of lemma 6.7. ⊣

The following proposition demonstrates how the quantifiers behave in mq-models. The classic quantifier rules are preserved for the worlds under the empty restriction. For the wurs, however, principles reminiscent of free interpretations of the quantifiers hold.

Proposition 6.14 For all $L \in \mathcal{K}^*$ and for all mq-models $\mathcal{M} = \langle\mathcal{W},\mathcal{R}\rangle$:

(i) $\not\models\ \forall x\varphi \to (t/x)\varphi$

(ii) If $x \notin RT(t)$, t free for x in $\varphi$and ϖ is pervasive for t w.r.t. x, then: $\mathcal{M},\varpi \models (\forall x\varphi \wedge \exists x\, x \equiv t) \to (t/x)\varphi$

(iii) If $x \notin RT(t)$:$\models \exists x\, x \equiv t$

(iv) $\models \forall x\varphi \to (t/x)\varphi$ $\qquad$ t free for x in φ

(v) If $\mathcal{M} \models \varphi$ then $\mathcal{M} \models \forall x\varphi$

Proof: Consider an arbitrary $L \in \mathcal{K}^*$:

(i) A countermodel can be constructed along the following lines. Define the qm-model $\mathcal{M} = \langle\mathcal{W},\mathcal{R}\rangle$ such that:

- $\mathcal{W} = w_1, w_2$ with:
  - $w_1 = \langle\{d_1\}, [\![\cdot]\!]_{w_1}, g\rangle$ with $[\![\cdot]\!]_{w_1}$ such that $d_1 = [\![a]\!]_{w_1} \in [\![P]\!]_{w_1}$.
  - $w_2 = \langle\{d_2\}, [\![\cdot]\!]_{w_2}, h\rangle$ with $[\![a]\!]_{w_2} = d_2 \neq d_1$.
- $\mathcal{R} = \{\langle w_2, w_1\rangle\}$

$\mathcal{M}$ is a proper mq-model since the domains of both w_1 and w_2 are singletons and, as such, for each there is only one suitable variable assignment. It can easily be verified that both $\mathcal{M}, w_1[\{a\}/w_2[\text{-}]] \models \forall xPx$ and $\mathcal{M}, w_1[\{a\}/w_2[\text{-}]] \not\models Pa$ ($\approx (a/x)Px$).

(ii) (sketch) Assume, for some arbitrary mq-model $\mathcal{M} = \langle\mathcal{W},\mathcal{R}\rangle$ and $\langle\mathcal{A},g\rangle[\varrho] \in \textsc{wur}(\mathcal{W})$ and some equally arbitrary $\varphi \in \textsc{Form}_L$, $x \in \textsc{Var}_L$ and $t \in \textsc{Term}_L$ such that $x \notin RT(\varphi)$, that $\mathcal{M},\langle\mathcal{A},g\rangle[\varrho] \models \forall x\varphi \wedge \exists x\, x \equiv t$. Due to the second conjunct there should be some $h \in \textsc{Ass}_{\mathcal{A}}$ such that $\mathcal{M},\langle\mathcal{A},h\rangle[\{\breve{\bar{x}}\}/\langle\mathcal{A},g\rangle[\varrho]] \models x \equiv t$. Consider this h. Since $\mathcal{M},\langle\mathcal{A},g\rangle[\varrho] \models \forall x\varphi$, also $\mathcal{M},\langle\mathcal{A},h\rangle[\{\breve{\bar{x}}\}/\langle\mathcal{A},g\rangle[\varrho]] \models \varphi$. By lemma 6.13 both $\mathcal{M},\langle\mathcal{A},g[x/h(x)]\rangle[\varrho] \models x \equiv t$ and $\mathcal{M},\langle\mathcal{A},g[x/h(x)]\rangle[\varrho] \models \varphi$. $\langle\mathcal{A},g\rangle[\varrho]$ could be assumed to be pervasive for x with respect to t. It only takes an easy check to verify that $\langle\mathcal{A},g[x/h(x)]\rangle[\varrho]$ is likewise pervasive. Since, moreover, x is free for t in φ we may conclude by an argument similar to the proof lemma 3.43 on page 45 above, that $\mathcal{M},\langle\mathcal{A},g[x/h(x)]\rangle[\varrho] \models (t/x)\varphi$. Because $x \notin RT(t)$ and x is free for t in φ, $x \notin RT((t/x)\varphi)$, $\langle\mathcal{A},g[x/h(x))[\varrho] \sim_{RT((t/x)\varphi)} \langle\mathcal{A},g\rangle[\varrho]$, i. e. $\langle\mathcal{A},g[x/h(x)]\rangle[\varrho]$ and $\langle\mathcal{A},g\rangle[\varrho]$ coincide on the interpretation of $RT((t/x)\varphi)$. This is sufficient to conclude that $\mathcal{M},\langle\mathcal{A},g\rangle[\varrho] \models (t/x)\varphi$.

(*iii*) Consider an arbitrary $t \in \textsc{Term}_L$ such that $x \notin RT(t)$, i. e. $t \in \{\overline{\breve{x}}\}$. Consider also arbitrary mq-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and an equally arbitrary $w = \langle \mathcal{A}, g \rangle \in \mathcal{W}$. It is then sufficient to prove that there is an $h \in \textsc{Ass}_\mathcal{A}$ such that $\llbracket x \rrbracket_{\langle \mathcal{A},h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\text{-}]]} = \llbracket t \rrbracket_{\langle \mathcal{A},h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\text{-}]]}$. Set $h = g[x/\llbracket t \rrbracket_{\langle \mathcal{A},g\rangle[\text{-}]}]$. Since $\llbracket t \rrbracket_{w[\text{-}]} \in D_w$, $g[x/\llbracket t \rrbracket_{\langle \mathcal{A},g\rangle[\text{-}]}] \in \textsc{Ass}_\mathcal{A}$. The following equalities now clinch the proof:

$$\llbracket x \rrbracket_{\langle \mathcal{A},g[x/\llbracket t \rrbracket_{\langle \mathcal{A},g\rangle[\text{-}]}]\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\text{-}]]} =_{x \notin \{\overline{\breve{x}}\}} \llbracket x \rrbracket_{\langle \mathcal{A},g[x/\llbracket t \rrbracket_{\langle \mathcal{A},g\rangle[\text{-}]}]\rangle} =$$
$$g[x/\llbracket t \rrbracket_{\langle \mathcal{A},g\rangle[\text{-}]}](x) = \llbracket t \rrbracket_{\langle \mathcal{A},g\rangle[\text{-}]} =_{t \in \{\overline{\breve{x}}\}} \llbracket t \rrbracket_{\langle \mathcal{A},g[x/\llbracket t \rrbracket_{\langle \mathcal{A},g\rangle[\text{-}]}]\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\text{-}]]}.$$

(*iv*) Directly from (*i*) and (*iii*) above.

(*v*) Suppose for contraposition $\not\models \forall x\varphi$. Then there must be some mq-model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and some $\langle \mathcal{A}, g \rangle$ such that $\mathcal{M}, \langle \mathcal{A}, g \rangle[\text{-}] \not\models \forall x\varphi$. Hence there must also be some $h \in \textsc{Ass}_\mathcal{A}$ such that $\mathcal{M}, \langle \mathcal{A}, h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\text{-}]] \not\models \varphi$. Finally, by lemma 6.13: $\mathcal{M}, \langle \mathcal{A}, g[x/h(x)]\rangle[\text{-}] \not\models \varphi$, and we are done. $\dashv$

6.3 λ-Abstraction

Corresponding to the modal notion of quantification as presented in the previous section, a notion of class abstraction could be introduced to the object language.[3] Let for each $x \in \textsc{Var}_L$ and each $\varphi \in \textsc{Form}_L$, $\hat{x} \cdot \varphi$ be a complex predicate symbol that in each wur $w[\varrho]$ is interpreted as that property which holds of all values for x in D_w for which φ is true in $w[\varrho]$.

Definition 6.15 For all $L \in \mathcal{K}^*$, $x \in \textsc{Var}_L$, $t \in \textsc{Term}_L$, $\varphi \in \textsc{Form}_L$ and all mq-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and all $\langle \mathcal{A}, g \rangle[\varrho] \in \textsc{Wur}(\mathcal{W})$:

(*i*) $\llbracket \hat{x} \cdot \varphi \rrbracket_{\langle \mathcal{A},g\rangle[\varrho]} :=$
$\{\llbracket x \rrbracket_{\langle \mathcal{A},h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\varrho]]} \mid h \in \textsc{Ass}_\mathcal{A} \;\&\; \mathcal{M}, \langle \mathcal{A}, h\rangle[\{\overline{\breve{x}}\}/\langle \mathcal{A},g\rangle[\varrho]] \models \varphi\}$

(*ii*) $\mathcal{M}, \varpi \models (\hat{x} \cdot \varphi)(t) :\iff \llbracket t \rrbracket_\varpi \in \llbracket \hat{x} \cdot \varphi \rrbracket_\varpi$

The semantical behaviour of $\hat{x} \cdot \varphi$ is much as expected:

Proposition 6.16 For all $L \in \mathcal{K}^*$, $x \in \textsc{Var}_L$, $t \in \textsc{Term}_L$, $\varphi \in \textsc{Form}_L$ and all mq-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and all $\langle \mathcal{A}, g \rangle[\varrho] \in \textsc{Wur}(\mathcal{W})$,
if $x \notin RT(t)$ and t free for x in φ and ϖ is pervasive for t w.r.t. x, then:
$\mathcal{M}, \varpi \models (\hat{x} \cdot \varphi)(t) \leftrightarrow (t/x)\varphi$

Proof: Analogous to the proof of proposition 6.14(*ii*), above. $\dashv$

The treatment of quantifiers in the previous section made that the domain of quantification in each wur $w[\varrho]$ coincided with D_w. Similarly, $\llbracket \hat{x} \cdot \varphi \rrbracket_{w[\varrho]} \subseteq D_w$. One could, however, wish one's notion of abstraction to cater for all substitution instances in each wur $w[\varrho]$, and not only for those that are guaranteed to be interpreted in D_w. Note that, e .g. $w[\varrho] \models Pt \vee \neg Pt (\approx (t/x)(Px \vee \neg Px))$ even if $\llbracket t \rrbracket_{w[\varrho]} \notin D_w$. At first sight it might seem that setting $\textsc{Ass}_{w[\varrho]}$ as the set of

[3] The results of this section are of an impressionistic character. Proofs, moreover, are largely omitted.

assignment functions that yields an object $d \in D_{w[\varrho]}$ (cf. definition 3.10 on page 28, above), rather than $d \in D_w$, as the value of each variable and subsequently defining abstraction as:
$[\![\hat{x} \cdot \varphi]\!]_{\langle \mathcal{A},g\rangle[\varrho]} :=$
$\{[\![x]\!]_{\langle \mathcal{A},h\rangle[\{\overline{x}\}/\langle \mathcal{A},\}\rangle[\varrho]]} \mid h \in \text{Ass}_{\langle \mathcal{A},g\rangle[\varrho]} \;\&\; \mathcal{M}, \langle \mathcal{A},h\rangle[\{\overline{x}\}/\langle \mathcal{A},\}\rangle[\varrho]] \models \varphi\}$,
would do the trick. The problem is, however, that $\langle \mathcal{A},h\rangle$ is not in general a feasible world of an mq-model, since $\text{Ass}_{\langle \mathcal{A},h\rangle[\varrho]}$ that are *not* suitable for $\mathcal{A}$.

A slightly different approach is called for. Things can be repaired by adopting the following definitions:

Definition 6.17 Define for each $\varpi = \varpi \in \text{WUR}$, $\varpi[\{x\}/w']$ by induction on ϖ as follows:

- $w[\text{-}][\{x\}/w'] := w[\{x\}/w']$
- $w[Z/\varpi'][\{x\}/w'] := w[Z \cup \{x\}/\varpi'[\{x\}/w']]$[4]

Definition 6.18 (The Interpretation of λ-terms) For all $L \in \mathcal{K}^*$, $x \in \text{Var}_L$, $t \in \text{Term}_L$, $\varphi \in \text{Form}_L$ and all mq-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$ and all $w[\varrho] \in \text{Wur}(\mathcal{W})$:

(*i*) $[\![\lambda x.\varphi]\!]_{w[\varrho]} = \{[\![x]\!]_{w'} \mid w' \in \mathcal{W} \;\&\; \mathcal{M}, w[\varrho][\{x\}/w'] \models \varphi\}$

(*ii*) $\mathcal{M}, \varpi \models (\lambda x.\varphi)(t) \iff [\![t]\!]_\varpi \in [\![\lambda x.\varphi]\!]_\varpi$

The semantical interpretation of $\lambda x.\varphi$ in a wur $w[\varrho]$ differs from that of $\hat{x} \cdot \varphi$ in the same wur in that the the former may contain objects outside D_w. It turns out that this restores the classical substitution principles for λ-terms in **PLuM**: This restores the classical substitution-rules, as witness:

Proposition 6.19 For all $L \in \mathcal{K}^*$, $x \in \text{Var}_L$, $t \in \text{Term}_L$, $\varphi \in \text{Form}_L$ and all mq-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$ and all $\langle \mathcal{A}, g\rangle[\varrho] \in \text{Wur}(\mathcal{W})$,
if t free for x in φ and ϖ is pervasive for t w.r.t. x, then:

$$\mathcal{M}, \varpi \models (\lambda x.\varphi)(t) \leftrightarrow (t/x)\varphi$$

Proof: By a combination of similar though more complex proofs as those of lemma 3.43 and proposition 6.14(*i*), above. ⊣

On the account of the previous section the classical principles of quantification can be regained for the wurs by adopting a universal domain for all worlds in the mq-model. The present treatment of λ-abstraction, however, facilitates an interpretation of the quantifiers such that they range over $D_{w[\varrho]}$ in each wur $w[\varrho]$ of the model:

Definition 6.20 For all $L \in \mathcal{K}^*$, $x \in \text{Var}_L$, $t \in \text{Term}_L$, $\varphi \in \text{Form}_L$ and all mq-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}\rangle$ and all $\langle \mathcal{A}, g\rangle[\varrho] \in \text{Wur}(\mathcal{W})$:

$$\mathcal{M}, \varpi \models \bigwedge x\varphi :\iff D_\varpi \subseteq [\![\lambda x.\varphi]\!]_\varpi{}^5$$

The range of any quantifier $\bigwedge x$ is then constituted by all substitution instances of x in a wur. This also reinstates the classical substitution rules for $\bigwedge x$:

[4] Note that *not* in general $\varpi[\{x\}/w] \in \text{WUR}(\mathcal{W})$.

[5] In a similar vein, it would seem, one could redefine the semantical clause for $\forall x\varphi$ as:
$\mathcal{M}, w[\varrho] \models \forall x\varphi :\iff D_w \subseteq [\![\lambda x.\varphi]\!]_{w[\varrho]}$

Proposition 6.21 For all $L \in \mathcal{K}^*$, $x \in \textsc{Var}_L$, $t \in \textsc{Term}_L$, $\varphi \in \textsc{Form}_L$ and all mq-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$ and $\varpi \in \textsc{wur}(W)$:
if t free for x in φ and ϖ is pervasive for t w.r.t. x then:
$\Vdash \bigwedge x\varphi \to (t/x)\varphi$

Proof: Analogous to the proof of proposition 6.19. $\dashv$

This observation concludes this section. Still the treatment of quantification in **PLuM**, I am sure, is far from complete. I have contented myself with merely outlining the general picture. For one thing, throughout this chapter it has tacitly been assumed that atomic formula $R(t_1, \ldots, t_n)$ are false in a wur $w[\varrho]$ if for some t_i, with $1 \leq i \leq n$, $[\![t]\!]_{w[\varrho]} \notin D_w$. Other choices could be made in this respect, which might result in that other logical principles obtain. I daresay, however, that the choices in this respect are largely dependent upon philosophical considerations. Neither do I think that these philosophical issues are more congenial to **PLuM** than they are to other quantified modal logics. Neither do I like to think about **PLuM** as providing important new insights in this respect. The philosophical justification of **PLuM** should be sought elsewhere, e. g. in its behaviour in face modal collapse threatening the intelligibility of modal predicate logic and the conceptual similarities of its semantics of *de re* readings of terms to Kaplan's notion of direct reference. Nothing has been said, moreover, about such matters as completeness in quantified **PLuM**. Especially with respect to the 'lighter' modal interpretations of the quantifiers I imagine some interesting results might be obtained.

6.4 Quine's Paradox Revisited

In his early papers on modality and intensionality Quine launched a powerful attack on quantified modal logic by pointing out that transparency of modal operators makes for the collapse of modal distinctions in the sense that $\varphi \to \Box\varphi$ would become derivable. With $\Box\varphi \to \varphi$ as an obvious principle for any concrete modal concept, this would render the truth of a formula equivalent to that formula being necessarily true.

Semantical investigations in the field of quantified modal logic have established that quantifying into modal contexts is possible without modal distinctions being annihilated. Føllesdal observed that this is achieved through distinguishing between general and singular terms. Modal operators are presumed to be transparent with respect to the singular terms, among which variables and presumably the individual constants. This in contradistinction to the general terms, with respect to which the modal operators are thought to be in general opaque. The substitution of general terms when occurring in a modal contexts is thus rendered impossible. On this account, Quine's argument as it was presented in the introduction of this thesis on page 6 above, founders because it ultimately depends on the feasibility of substitution of $\imath$-terms, which would rank among the general terms, in intensional contexts. Stipulating modal operators to be opaque for general terms, however, leaves Quine's argument unscathed as an argument against fully transparent modal contexts.

Two fundamental thoughts underlie this thesis. The first is that there are essentially predicate logical principles of modality apart from the interaction of modality and quantification. For these any modal predicate logic should account. The difference between *de dicto* and *de re* readings of terms in modal contexts are taken to be evidence hereof. The second concerns the idea to conceive of modal operators as binding terms in their scope. These views precipitated a syntactical revision of the formal languages for modal predicate logic. The modal operators got indexed by a set of terms they were presumed not to bind. The unquantified modal logic **PLuM** provides a semantics for such languages, in which the terms that are not in the index of the modal operator obtain a *de dicto* reading; the others will get a *de re* interpretation.

As things turned out, the only restriction on the substitutivity of (occurrences of) terms is that they are not bound by a modal operator. No distinction between general terms and singular terms had to be imposed in this respect. Still, no modal collapse is immanent in **PLuM**. This time, however, because replacement and not substitutivity, is compromised to some extent. As it is presented in this thesis, the validity of Quine's argument depends on the eligibility of replacing $(\imath x \cdot x \equiv a) \equiv (\imath x \cdot x \equiv a \wedge \varphi)$ by φ within the scope of a modal operator. It can rightly be observed that in any logic dealing with $\imath$-terms $(\imath x \cdot x \equiv a) \equiv (\imath x \cdot x \equiv a \wedge \varphi)$ and φ should be equivalent in some sense. This equivalence would also sanction their replacement. So far the formal interpretation of $\imath$-operators in **PLuM** has not been dealt with in this thesis. $\imath$-terms could be added to the languages of $\mathcal{L}$ or $\mathcal{K}$, and some semantical clause interpreting them along the following lines would seem to be in order:

Definition 6.22 Define for each $L \in \mathcal{K}$, all $x \in \textsc{Var}_L$ and $\varphi \in \textsc{Form}_L$, for all mq-models $\mathcal{M} = \langle \mathcal{W}, \mathcal{R} \rangle$, and all $w[\varrho] \in \textsc{wur}(\mathcal{W})$, $[\![\imath x \cdot \varphi]\!]_{w[\varrho]}$ inductively as:

- $[\![\imath x \cdot \varphi]\!]_{w[\text{-}]} := \begin{cases} d & \text{if } d \in [\![\hat{x} \cdot \varphi]\!]_{w[\text{-}]} \text{ and } |[\![\hat{x} \cdot \varphi]\!]_{w[\text{-}]}| = 1 \\ \uparrow & \text{otherwise} \end{cases}$
- $[\![\imath x \cdot \varphi]\!]_{w[Z/\varpi]} := \begin{cases} d & \text{if } d \in [\![\hat{x} \cdot \varphi]\!]_{w[Z/\varpi]},\ |[\![\hat{x} \cdot \varphi]\!]_{w[Z/\varpi]}| = 1 \\ & \text{and } \imath x \cdot \varphi \notin Z \\ d' & \text{if } d' \in [\![\lambda x \cdot \varphi]\!]_{w[Z/\varpi]},\ |[\![\hat{x} \cdot \varphi]\!]_{w[\text{-}]}| = 0, \\ & |[\![\lambda x \cdot \varphi]\!]_{w[Z/\varpi]}| = 1 \text{ and } \imath x \cdot \varphi \notin Z \\ [\![\imath x \cdot \varphi]\!]_{\varpi} & \text{if } \imath x \cdot \varphi \in Z \\ \uparrow & \text{otherwise} \end{cases}$

N.B.: $|Z|$ denotes the cardinality of Z and for each term t, $[\![t]\!]_{\varpi} = \uparrow$ iff $[\![t]\!]_{\varpi}$ is undefined.

Note that this would compel one to revise the elementary semantics of **PLuM** to some extend. Undefined terms have now also to be accounted for. Formally, it may be quite convenient to treat undefined terms as defined terms that are not in the extension of any relation symbol. Still, philosophical considerations may make one decide otherwise. Whatever particular choices are made, the above definition certainly yields $\models ((\imath x \cdot x \equiv a) \equiv (\imath x \cdot x \equiv a \wedge \varphi)) \leftrightarrow \varphi$. Since for each φ and each x, $\imath x \cdot \varphi$ is presumed to be a genuine term and as such it may occur in the indices, however, $\models ((\imath x \cdot x \equiv a) \equiv (\imath x \cdot x \equiv a \wedge \varphi)) \leftrightarrow \varphi$ does *not* hold in general as the following example illustrates:

Example 6.23 Let $\mathcal{M} = \langle \{w_1 w_2\}, \mathcal{R} \rangle$ with $w_1 \mathcal{R} w_2$ and, $w_1 = \langle \{d_1\}, [\![\cdot]\!]_{w_1}, g \rangle$ and $w_2 = \langle \{d_2\}, [\![\cdot]\!]_{w_2} \rangle$ with for some propositional variable p, $[\![p]\!]_{w_1} = \{\langle\rangle\}$ and

$[\![p]\!]_{w_2} = \emptyset$. As can easily be verified $\mathcal{M}, w_2[\{\imath x \cdot (x \equiv a \wedge p), \imath x \cdot x \equiv a\}/w_1[\text{-}]] \not\models p$. Since d_1 is the only element in D_{w_1}, and so $[\![a]\!]_{w_1} = d_1$, the following equations, however, hold as well:

$$[\![\imath x \cdot x \equiv a]\!]_{w_2[\{\imath x \cdot (x \equiv a \wedge p), \imath x \cdot a \equiv x\}/w_1[\text{-}]]} = [\![\imath x \cdot x \equiv a]\!]_{w_1[\text{-}]} = d_1 = [\![\imath x \cdot x \equiv a \wedge p]\!]_{w_1[\text{-}]} = [\![\imath x \cdot x \equiv a \wedge p]\!]_{w_2[\{\imath x \cdot (x \equiv a \wedge p), \imath x \cdot x \equiv a\}/w_1[\text{-}]]}.$$

Consequently:

$$\mathcal{M}, w_2[\{\imath x \cdot (x \equiv a \wedge p), \imath x \cdot x \equiv a\}/w_1[\text{-}]] \not\models \imath x \cdot (x \equiv a \wedge p) \equiv \imath x \cdot x \equiv a \leftrightarrow p.$$

Replacement of one formula for another within the scope of a modal operator, however, requires equivalence in all wurs of at least depth 1, as was demonstrated in section 3.2.4, above. This failing to hold for $(\imath x \cdot x \equiv a) \equiv \imath x \cdot (x \equiv a \wedge \varphi)$ and φ, makes that Quine's argument breaks down in **PLuM**. Entirely transparent modal operators, however, can be made perfect sense of in **PLuM**, by indexing them by the set of all terms. Still, different possibilities can be accounted for.

Bibliography

Bencivenga, E. [1986], 'Free Logics', in: D. Gabbay & F. Guenther (eds.), *Handbook of Philosophical Logic*, Vol. III, D. Reidel, Dordrecht, 1986

Bencivenga, E. [1991], 'Free Semantics', in: K. Lambert (ed.), *Philosophical Applications of Free Logic*, Oxford U.P., New York, 1991

Benthem, J. F. A. K. van [1996], *Exploring Logical Dynamics*, Chapt. 9, Studies in Logic, Language and Information, CSLI, Stanford, Ca., 1996

Beth, E. W. [1962], 'Extension and Intension', in: *Logic and Language. Studies Dedicated to Professor Rudolf Carnap on the Occasion of his Seventieth Birthday*, Synthese Library, D. Reidel, Dordrecht, 1962

Carnap, R. [1948], *Meaning and Necessity. A Study in Semantics and Modal Logic*, Chicago U.P., Chicago, 1948, 2nd impression

Cartwright, R. [1971], 'Identity and Substitutivity', in: *Philosophical Papers*, The MIT Press, Cambridge, Mass., 1987

Cartwright, R. [1979], 'Indicernibility Principles', in: *Philosophical Papers*, The MIT Press, Cambridge, Mass., 1987

Chellas, B. F. [1980], *Modal Logic. An Introduction*, Cambridge U.P., Cambridge, 1980

Chisholm, R. M. [1967], 'Identity through Possible Worlds: Some Questions', *Noûs* **1** (1967), 1–8

Church, A. [1951], 'A Formulation of the Logic of Sense and Denotation', in: P. Henle, H. M. Kallen & S. K. Langer (eds.), *Structure, Method and Meaning. Essays in the Honor of Henry M. Sheffer*, The Liberal Arts Press, New York, 1951

Church, A. [1982], 'A Remark Concerning Quine's Paradox of Modality', in: N. Salmon & S. Soames (eds.), *Propositions and Attitudes*, Oxford Readings in Philososphy, Oxford U.P., Oxford, 1988

Cocchiarella, N. [1984], 'Philosophical Perspectives on Quantification in Tense and Modal Logic', in: D. Gabbay & F. Guenther (eds.), *Handbook of Philosophical Logic*, Vol. II, D. Reidel, Dordrecht, 1984

Davidson, D. & Hintikka, J. (eds.) [1969], *Words and Objections. Essays on the Work of W.V. Quine* , Synthese Library 21, D. Reidel, Dordrecht, 1969

Does J. van der [1992], *Extensionele Logica, syllabus 1992/1993*, ILLC/FdW, University of Amsterdam, Amsterdam, 1992

Donnellan, K. S. [1971], 'Reference and Definite Descriptions', in: S. P. Schwartz (ed.), *Naming, Necessity and Natural Kinds*, pp 42–65, Cornell U.P., Ithaca, 1977

Fitch, F. B. [1949], 'The Problem of the Morningstar and the Eveningstar', in: I. M. Copi & J. A. Gould (eds.), *Contemporary Readings in Logical Theory*, Macmillan, New York, 1967

Føllesdal, D. [1969], 'Quine on Modality', in: D. Davidson & J. Hintikka (eds.), *Words and Objections* , Synthese Library 21, D. Reidel, Dordrecht, 1969

Føllesdal, D. [1986], 'Essentialism and Reference', in: L. E. Hahn & P. A. Schilpp (eds.), *The Philosophy of W.V. Quine*, Vol. XVIII of *The Library of Living Philosophers*, Open Court, La Salle, Ill., 1986

Frege, G. [1892a], 'Funktion und Begriff', in: I. Angelelli (ed.), *Kleine Schriften*, Georg Olms, Hildesheim, 1967

Frege, G. [1892b], 'Über Sinn und Bedeutung', in: I. Angelelli (ed.), *Kleine Schriften*, Georg Olms, Hildesheim, 1967

Gamut, L. T. F. [1991], *Logic, Language, and Meaning*, Vol. 2: Intensional Logic and Logical Grammar, Chicago U.P., Chicago, 1991

Garson, J. W. [1984], 'Quantification in Modal Logic', in: D. Gabbay & F. Guenther (eds.), *Handbook of Philosophical Logic*, Vol. II, D. Reidel, Dordrecht, 1984

Garson, J. W. [1991], 'Applications of Free Logic to Quantified Intensional Logic', in: K. Lambert (ed.), *Philosophical Applications of Free Logic*, Oxford U.P., New York, 1991

Geach, P. T. [1968], *A History of the Corruptions of Logic. An Inaugural Lecture*, Leeds U.P., Leeds, 1968

Gerbrandy, J. [1997], 'Questions of Identity', in: P. Dekker, M. Stokhof & Y. Venema (eds.), *Proceedings of the 11th Amsterdam Colloquium*, pp 133–138, ILLC, Amsterdam, 1997

Grayling, A. C. [1982], *An Introduction to Philosophical Logic*, The Harvester Press, Brighton, 1982

Groenendijk, J., Stokhof, M. & Veltman, F. [1996], 'Coreference adn Modality', in: S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, Basil Blackwell, Oxford, 1996

Haack, S. [1978], *Philosophy of Logics*, Cambridge U.P., Cambridge, 1978

Hughes, G. E. & Cresswell, M. J. [1968], *An Introduction to Modal Logic*, Methuen, London, 1974

Jongh, D. de & Veltman, F. [1995], *Intensional Logics*, FdW, University of Amsterdam, Amsterdam, 1995

Kaplan, D. [1969], 'Quantifying In', in: D. Davidson & J. Hintikka (eds.), *Words and Objections*, Vol. 21 of *Synthese Library*, D. Reidel, Dordrecht, 1969

Kaplan, D. [1973], 'Bob and Carol and Ted and Alice', in: J. H. et.al. (ed.), *Approaches to Natural Language*, D. Reidel, Dordrecht, 1973

Kaplan, D. [1975], 'How to Russell a Frege-Church', in: M. J. Loux (ed.), *The Possible and the Actual*, Cornell U.P., Ithaca, 1979

Kaplan, D. [1977], 'Demonstratives', in: J. Almog, J. Perry & H. Wettstein (eds.), *Themes from Kaplan*, Oxford U.P., New York, 1989

Kaplan, D. [1978a], 'Dthat', in: P. A. French, T. E. Ueling, Jr & H. K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, Minnesota U.P., Minneapolis, 1979

Kaplan, D. [1978b], 'On the Logic of Demonstratives', in: P. A. French, T. E. Ueling, Jr & H. K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, Minnesota U.P., Minneapolis, 1979

Kaplan, D. [1979], 'Transworld Heir Lines', in: M. J. Loux (ed.), *The Possible and the Actual*, Cornell U.P., Ithaca, 1979

Kaplan, D. [1985], 'Opacity', in: L. E. Hahn & P. A. Schilpp (eds.), *The Philosophy of W. V. Quine*, Vol. XVIII of *The Library of Living Philosophers*, Open Court, La Salle, Ill., 1986

Kaplan, D. [1989], 'Afterthoughts', in: J. Almog, J. Perry & H. Wettstein (eds.), *Themes from Kaplan*, Oxford U.P., New York, 1989

Kneale, W. [1962], 'Modality De Dicto and De Re', in: E. Nagel, P. Suppes & A. Tarski (eds.), *Logic, Methodology and Philosophy of Science. Proceedings of the 1960 International Congress*, Stanford U.P., Stanford, 1962

Kripke, S. A. [1963], 'Semantical Considerations on Modal Logic', in: *Reference and Modality*, Oxford Readings in Philosophy, Oxford U.P., Oxford, 1971

Kripke, S. A. [1971], 'Identity and Necessity', in: S. P. Schwartz (ed.), *Naming, Necessity and Natural Kinds*, Cornell U.P., Ithaca, 1977

Kripke, S. A. [1972], 'Naming and Necessity', in: G. Harman & D. Davidson (eds.), *Semantics of Natural Language*, D.Reidel, Dordrecht, 1972

Kripke, S. A. [1979], 'A Puzzle about Belief', in: A. Margalit (ed.), *Meaning and Use. Papers Presented at the Second Jerusalem Philosophical Encounter April 1976* , Synthese Language Library 3, D. Reidel, Ithaca, 1977

Kripke, S. A. [1980], *Naming and Necessity*, Basil Blackwell, Oxford, 1980

Lemmon, E. J. [1963], 'A Theory of Attributes Based on Modal Logic', in: *Proceedings of a Colloquium on Modal and Many-Valued Logics, Helsinki, 23–29 August, 1962*, *Acta Philosphica Fennica* **XVI** (1963)

Lemmon, E. J. [1977], *An Introduction to Modal Logic. The "Lemmon Notes"* , American Philosophical Quarterly Monograph Series 11, Basil Blackwell, Oxford, 1977

Lewis, D. [1968], 'Counterpart Theory and Quantified Modal Logic', in: M. J. Loux (ed.), *The Possible and the Actual*, Cornell U.P., Ithaca, 1979

Lewis, D. [1973], 'Possible Worlds', in: M. J. Loux (ed.), *The Possible and the Actual*, Cornell U.P., Ithaca, 1979

Linsky, L. [1969], 'Reference, Essentialism, and Modality', in: L. Linsky (ed.), *Reference and Modality*, Oxford Readings in Philosophy, Oxford U.P., Oxford, 1971

Linsky, L. (ed.) [1971], *Reference and Modality*, Oxford Readings in Philosophy, Oxford U.P., Oxford, 1971

Loux, M. J. (ed.) [1979], *The Possible and the Actual*, Cornell U.P., Ithaca, 1979

Łukasiewicz, J. [1953], 'A System of Modal Logic', in: *Actes du XIème Congrès international de philosophie*, Vol. XIV: Volume complémentaire et communications du colloque de logique, North-Holland, Amsterdam, 1953

Marcus, R. B. [1961], 'Modalities and Intensional Languages', in: *Modalities. Philosophical Essays*, Oxford U.P., New York, 1993

Marcus, R. B. [1967], 'Essentialism in Modal Logic', in: *Modalities. Philosophical Essays*, Oxford U.P., New York, 1993

McCulloch, G. [1989], *The Game of the Name*, Oxford U.P., Oxford, 1989

Montague, R. [1960], 'Logical Necessity, Physical Necessity, Ethics, and Quantifiers', in: *Formal Philosophy. Selected Papers of Richard Montague*, edited by R.H. Thomason, Yale U.P., New Haven, 1974

Parsons, T. [1969], 'Essentialism and Quantified Modal Logic', in: *Reference and Modality*, Oxford Readings in Philosophy, Oxford U.P., Oxford, 1971

Popkorn, S. [1994], *First Steps in Modal Logic*, Cambridge U.P., Cambridge, 1994

Quine, W. V. O. [1936], 'Truth by Convention', in: H. Feigl & W. Sellars (eds.), *Readings in Philosophical Analysis*, Appleton-Century-Crofts, New York, 1949

Quine, W. V. O. [1940], *Mathematical Logic*, W.W. Norton, New York, 1940

Quine, W. V. O. [1943], 'Notes on Existence and Necessity', *Journal of Philosophy* **40** (1943), 113–127

Quine, W. V. O. [1947], 'The Problem of Interpreting Modal Logic', in: I. M. Copi & J. A. Gould (eds.), *Contemporary Readings in Logical Theory*, MacMillan, New York, 1967

Quine, W. V. O. [1948], 'On What There Is', in: *From a Logical Point of View*, Harvard U.P., Cambridge, Mass., 1953

Quine, W. V. O. [1951], 'Two Dogmas of Empiricism', in: *From a Logical Point of View*, Harvard U.P., Cambridge, Mass., 1953

Quine, W. V. O. [1953a], *From a Logical Point of View*, Harvard U.P., Cambridge, Mass., 1953, first edition

Quine, W. V. O. [1953b], 'Reference and Modality', in: *From a Logical Point of View*, Harvard U.P., Cambridge,Mass., 1953/1961[6]

Quine, W. V. O. [1953c], 'Three Grades of Modal Involvement', in: *Actes du XIème Congrès international de philosophie*, Vol. XIV: Volume complémentaire et communications du colloque de logique, North-Holland, Amsterdam, 1953

Quine, W. V. O. [1956], 'Quantifiers and Propositional Attitudes', in: *Reference and Modality*, Oxford Readings in Philosophy, Oxford U.P., Oxford, 1971

Quine, W. V. O. [1960], *Word and Object*, The MIT Press, Cambridge, Mass., 1960

Quine, W. V. O. [1961], *From a Logical Point of View*, Harper & Row, New York, 1961, second edition

Quine, W. V. O. [1977], 'Intensions Revisited', in: P. A. French, T. E. Ueling, Jr & H. K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, Minnesota U.P., Minneapolis, 1979

Quine, W. V. O. [1986], *Philosophy of Logic*, Harvard U.P., Cambridge, Mass., 1986, second edition

Russell, B. [1905], 'On Denoting', in: *Logic and Knowledge. Essays 1901-1950*, edited by R.C. Marsh, George, Allen & Unwin, London, 1956

Salmon, N. U. [1981], *Reference and Essence*, Chapt. 1, Princeton U.P., Princeton, 1981

Schwartz, S. P. (ed.) [1977], *Naming, Necessity, and Natural Kinds*, Cornell U.P., Ithaca, 1977

Scott, D. [1970], 'Advice on Modal Logic', in: K. Lambert (ed.), *Philosophical Problems in Logic*, Synthese Library, D. Reidel, Dordrecht, 1970

Searle, J. R. [1958], 'Proper Names', in: P. F. Strawson (ed.), *Philosophical Logic*,

[6]The first and second edition of *From a Logical Point of View* differ substantially with respect to this paper. Reference will be made to the respective editions of the book when citing from this article.

Oxford Readings in Philosophy, Oxford U.P., Oxford, 1967

Smullyan, A. F. [1948], 'Modality and Description', in: L. Linsky (ed.), *Reference and Modality*, Oxford Readings in Philosophy, Oxford U.P., Oxford, 1971

Stalnaker, R. [1984], *Inquiry*, The MIT Press, Cambridge, Mass., 1984

Tarski, A. [1944], 'The Semantic Conception of Truth', *Philosophy and Phenomenological Research* **4** (1944), 341–375

Veltman, F. [1985], *Logics for Conditionals*, Ph.D. thesis, FdW UvA, Amsterdam, 1985

Wiggins, D. [1965], 'Identity-statements', in: R. J. Butler (ed.), *Analytical Philosophy - second series*, Basil Blackwell, Oxford, 1965

Wittgenstein, L. [1922], *Tractatus Logico-Philosophicus*, German-English edition, English translation by D.F. Pears & B.F. McGuinness, Routledge & Kegan Paul, London, 1971