

Identity and Modality

A Thesis Submitted to the College of
Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts
in the Department of Philosophy
University of Saskatchewan
Saskatoon

By

Kenneth Kin Hung Chung

Spring 2003

© Copyright Kenneth Kin Hung Chung, 2002. All rights reserved.

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying, publication, or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Philosophy
University of Saskatchewan
Saskatoon, Saskatchewan
S7N 5A5

Abstract

In 1970, Saul Kripke gave the watershed lectures that is *Naming and Necessity*. One of his claims, which had far reaching implications, was that all identities are necessary. Despite his wide influence, many worries remained. This essay explores two of the most popular worries with Kripke's project: Quine's objection of referential opacity and the problems of transworld identification. I hope to show that both of these worries are merely worries, and the air of serious objection they both carry can be dispelled.

Acknowledgements

My greatest debt in the composition of this thesis is to my supervisor, Karl Pfeifer. I wish to acknowledge his patience with earlier drafts of my work. His insightful comments and good philosophical sense have greatly shaped this essay. I am also grateful for the efforts of Sarah Hoffman and Anthony Jenkins who have read through various parts of the thesis and provided numerous helpful comments. I also wish to thank Eric Dayton for fruitful discussion of some philosophical matters. A part of an earlier version of this thesis was delivered as a talk for the Department of Philosophy at the University of Saskatchewan. The subsequent discussion forced me to rethink certain issues and rewrite certain sections. I wish to acknowledge the department for that, for the kindness and encouragement I have received here, and for letting me do this.

My dissatisfaction at various times with earlier versions of my work were occasionally accompanied by unpleasant dispositions. I wish to thank all my friends for putting up with it.

For my parents

Contents

1	INTRODUCTION	1
2	OPACITY	7
2.1	Quine's modal argument	9
2.1.1	De re and de dicto	10
2.2	On identity	15
2.2.1	The law of substitutivity	15
2.2.2	The principle of identity	16
2.3	Indiscernibility principles	20
2.3.1	Quine's claim of referential opacity	27
2.3.2	Purely referential	31
2.3.3	Quantification	34
2.4	On necessity	37
3	REFERENCE	39
3.1	Naive descriptive theory of names	40
3.2	Rigid designation	42
3.2.1	De facto and de jure	45
3.2.2	Obstinate and persistent	47
3.2.3	Wide scope	49
3.3	Transworld identity	52
3.3.1	Three problems	53
3.3.2	Some proposed solutions	56
3.3.3	By analogy	60
3.3.4	Some remarks	63
3.4	Possible worlds	68
4	CONCLUSION	71
4.1	Wreen's argument	71
4.2	Final remarks	80

Chapter 1

Introduction

In 1970, Saul Kripke famously argued against a longheld belief about the relation between necessity and a prioricity by demonstrating the existence of true statements that were necessary and *a posteriori*. Taking Frege's puzzle of Hesperus and Phosphorus, he argued that the statement, (α), "Hesperus is Phosphorus" is one such statement—necessary and a posteriori. To say that a statement is necessarily true roughly means to say that it must be true, and that it could not be false. To say a statement was a priori true, again roughly speaking, meant that we could know its truth by reflection alone without the need of empirical observations. So before Kripke, almost everyone thought that statements that were necessarily true were also a priori true, and vice versa. The thought behind this was simple enough: for how could something that we learn by reflection alone possibly be false? How could we learn that something must be necessarily true if we had to look out into the world to find it out? But since notions of a prioricity (and its counterpart notion of a posteriority) are epistemological notions and notions of necessity (and its counterpart notion of contingency) are metaphysical notions, any strict relation between them

will be tenuous. And so Kripke drove a wedge between the two notions and argued that they were independent. Thus, in addition to having necessary a priori truths and contingent a posteriori truths, Kripke presented and defended the existence of truths that were necessary a posteriori and some that were contingent a priori.

Now the existence of contingent a priori truths has produced odd results (Tharp's theorems), and it has been subjected to its own criticisms, but only because Kripke seems to extend what counts as a priori to the stipulations we often make. So the question of whether these contingent truths are a priori or a posteriori seems, to many philosophers, a taxonomic issue. But the existence of necessary a posteriori truths is sometimes taken to be one of the most important discoveries in recent analytic philosophy. And even now, over thirty years later, some philosophers are puzzled by it, some refuse to believe it, and some are still excited by it. Kripke however was not alone in making claims like these. Hilary Putnam famously argued for a similar relation between natural kinds, like water, and their terms like "water" and " H_2O "; and though the thought experiments Putnam used were similar to Kripke's deployment of the device of possible worlds, he drew different, but not contrary, conclusions. In any case, a large part of their arguments was a claim about language and the confusions we have with the meanings of words. But since I am mostly interested in the necessity of identity statements, let me say a little about what Kripke did.

To accomplish the necessity of an identity statement such as (α), Kripke had to do at least two things: (i) show that identities between objects are necessary, and (ii) show that proper names are rigid designators. The first seems an indisputable metaphysical fact about objects, for how could an object be anything but itself? Many of the objections raised against (i) are objections against useful or meaningful

talk of necessity and possibility. The second is a fact about language, but it seemed to carry with it many implausible metaphysical and epistemological assumptions, like transworld identification.

The history of modal logic might be said to begin with Leibniz who introduced the useful notion of a possible world. With this notion, he said that a statement is necessarily true just in case it is true in all possible worlds. A statement is possibly true just in case it is true in some possible world. And of course, a statement is contingently true just in case it is true in the actual world but false in some possible world. But not all the modal logicians following Leibniz used this device. Rudolf Carnap famously associated necessity with analyticity. And his friend, Willard Van Orman Quine, argued even more famously against both notions. Modal logicians such as Ruth Barcan Marcus and Arthur Smullyan took Quine to task for failing to observe distinctions. Quine's views however grew more sophisticated. Sentences, he said, with modal operators on them—such as “It is necessary that: 9 is greater than 7” or equivalently “ $\Box(9 > 7)$ ”—were referentially opaque, and making sense of such contexts can only lead to serious trouble.¹ It led to serious trouble because Leibniz's law told us that we can exchange co-referential terms *salva veritate* but when we exchanged “9” for “the number of planets” we arrived at a falsehood. In light of this difficulty, it makes some sense why Kripke made the claim about the necessity of identity only concerning proper names and not referential descriptions like “the number of planets”. But Kripke still needed a semantic argument for all of this, and not simply a formal stipulation that descriptions should be excluded from his theory. He thus came up with the claim that proper names are rigid designators.

¹The corresponding symbol for possibility is “ \Diamond ” so “It is possible that: 9 is greater than 7” would be “ $\Diamond(9 > 7)$ ”. And so “It is contingent that: 9 is greater than 7”, which is false, should be read as “ $9 > 7 \ \& \ \Diamond\neg(9 > 7)$ ”.

To explain that proper names are rigid designators, I must first explain the technical term of “rigid designator”. One common explanation of the term is that something (a name or a description) is a rigid designator if and only if it refers to the same object in all possible worlds. But this formulation is ambiguous, and Kripke explicitly warns against one interpretation. To say that a proper name is a rigid designator is *not* to say that we could not have used the name differently or that the object the name refers to could not have had a different name. Someone may have named his dog “Ludwig Wittgenstein” (John Searle, I believe, has). There are then instances when the use of the name “Ludwig Wittgenstein” refers to a dog, and not the Austrian philosopher. I also think it is possible that the Wittgensteins might have given Ludwig a different first name, so that Ludwig Wittgenstein might have been called “Charlie Wittgenstein”, as doubtful as it is. But those possibilities are not at issue with the topic of rigid designation. To say that a proper name is a rigid designator implies that the use of a name refers to that object in all possible worlds. So if I use a proper name to refer to x in the actual world, then my use of that proper name refers to x in all possible worlds. And so in general, proper names differ from descriptions. Now it makes sense to say “Wittgenstein may not have gone on to be the most celebrated student of Bertrand Russell”, if I were entertaining what other ventures Wittgenstein may have undertaken had he not studied philosophy; whereas it makes very *little* sense to say, “Wittgenstein may not have gone on to be Wittgenstein”. Being Wittgenstein is a necessary feature of Wittgenstein, while being the most celebrated student of Bertrand Russell is certainly only contingent. That is, the description “the most celebrated student of Bertrand Russell” comes apart from Wittgenstein.

If Kripke is right about proper names, then it should come as no surprise that

statements consisting of the identity sign flanked by proper names are necessarily true. Let us consider the following:

(α) Hesperus is Phosphorus.

This sentence expresses a true proposition. According to philosophical folklore, the term “Hesperus” was used to refer to a certain celestial body that appeared *in the night* at a certain time of the year; whereas “Phosphorus” was used to refer to a certain celestial body that appeared *in the morning* at a different time of the year. It was upon scientific discovery, that it was learned that the two names refer to the same celestial body, namely the planet Venus. Whenever we *use* the name “Hesperus” we are talking about the object Venus. Again, when we use the name “Phosphorus” we are still talking about Venus. So in saying “Hesperus is Phosphorus”, we are using those words to refer directly to the object. Since it is necessary that any object is identical to itself, the proposition expressed by (α) is necessarily true.

But in addition to the Quinean objections, there was the objection stemming from worries about Kripke’s free and easy use of possible worlds, especially in his characterization of rigid designation. Though the objection is often stated as a problem of transworld identification, there are basically two related problems. One was that it made little sense to say that an object existed in more than one possible world, and the second was that there was no criterion of identifying the same object in another world. So my thesis is an attempt at resolving the issues of referential opacity and transworld identification. I hope to show that the issue of referential opacity cannot make any substantial argument against quantified modal logic or against our intuitions about necessity and possibility. As for the problems of transworld identification, I hope to show that they result from a misleading picture of possible worlds and a misunderstanding of what counts as identification; and that therefore, there is

no difficulty at all about transworld identification.

Chapter 2

Opacity

Kripke essentially provides two arguments for the necessity of identity: one which I will call the semantic argument and the second the formal argument. The first is expressed by the following passage:

... it was clear from $(\forall x)\Box(x = x)$ and Leibnitz's law that identity is an "internal" relation: $(\forall x)(\forall y)(x = y \rightarrow \Box x = y)$. (What pairs (x, y) could be counterexamples? Not pairs of distinct objects, for then the antecedent is false; nor any pair of an object and itself, for then the consequent is true.)¹

Though this semantic argument has been difficult to refute, the formal argument, originating with Ruth Barcan Marcus, has been subject to many objections. I shall consider some of the more convincing challenges. From the impact that his lectures would later have, Kripke's presentation of Barcan's proof that all identities are necessary would become the one most often reproduced.² He begins with what he calls the "law of substitutivity".

¹Notation has been altered to conform with this essay. Kripke 1980. p. 3

²Kripke does not in fact present this proof in *Naming and Necessity*, but rather in the lecture that became "Identity and Necessity", Kripke 1971. Notation has been altered to conform with this essay.

$$(1) (\forall x)(\forall y)[(x = y) \rightarrow (Fx \rightarrow Fy)]$$

This principle generally states that if two things are identical, anything that can be said true of one thing, must be true of the other. And if it is true that any object is *necessarily* self-identical, then

$$(2) (\forall x)\Box(x = x).$$

By substitution of the predicate in (2) into F , we arrive at

$$(3) (\forall x)(\forall y)[(x = y) \rightarrow (\Box(x = x) \rightarrow \Box(x = y))].$$

From the truth of $\Box(x = x)$, that any object is necessarily self-identical, we may drop it to conclude

$$(4) (\forall x)(\forall y)[(x = y) \rightarrow \Box(x = y)]$$

If two objects are identical, then they are necessarily identical. Notwithstanding the motivations for contingent identities, many philosophers had serious worries about the proof. I wish to consider the more common worries. Many presentations of these worries often seem muddled and confused. They begin usually with Quine's argument that modal sentences are referentially opaque, and so quantification into modal contexts is illegitimate. But there have been replies which seem to miss the point of the objection. Questions concerning identity, indiscernibility principles, and objectual quantification are raised that give an awful and thick air of doubt around Kripke's claim. My aim is to clear the air a little, and determine whether the formal argument for the necessity of identity is well-grounded, or whether another is nearby.

2.1 Quine's modal argument

Quine's objections to quantified modal logic and to essentialism are often seen to undermine Kripke's program, despite the fact that most of them were articulated well before Kripke's work. People often cite Kripke's other arguments to answer these objections, but discussion of Quine's worries is difficult, as he makes many subtle points surrounding his more obvious ones. Moreover, it is not obvious at first glance exactly how Quine's objections raise a problem for Kripke, or even if they do at all. Nonetheless, I shall begin with his famous example and try to see what he might be saying. Consider the following:

(5) the number of planets = 9

(6) 9 is necessarily greater than 7

By substitution, we arrive at

(7) the number of planets is necessarily greater than 7³

Now if we read (6) as "it is necessary that nine is greater than seven", we should then read (7) as "it is necessary that the number of planets is greater than seven". And hence, (7) should be read as false: it is not necessary that the number of planets is greater than seven. Substitution of "the number of planets" for "9" must be legitimate on formal grounds; and it should at least be allowed if it is allowed in Kripke's proof for the necessity of identity. Perhaps the conclusion we are meant to draw is that sentences with modal operators do not suggest essential properties; in particular, "— is necessarily greater than 7" does not point to a property some object can have. This is a reasonable assumption, given Quine's own distaste for essentialism and his refusal to separate the issues.

³This is a variation of an argument in Quine 1980, pp. 139–159.

2.1.1 De re and de dicto

Hence, the modal logicians claim there to be a distinction that Quine does not acknowledge in his argument. There is, they say, a difference between *de re* and *de dicto* modal statements.⁴ Roughly speaking, the former are predicated of objects, whereas the latter are predicated of propositions or sentences about those objects. The sentences (6) and (7) are then ambiguous. The distinction can be defended in the following way:

How shall we read (6)? We have two options for interpreting its truth conditions.

(6a) “9 is necessarily greater than 7” is true *if and only if* “9 is greater than 7” is necessarily true.⁵

(6b) “9 is necessarily greater than 7” is true *if and only if* “there is a unique number 9, which is necessarily greater than 7” is true.

The difference between the two seems trivial here; it is not obvious how the two truth conditions, expressed on the right hand side of both (6a) and (6b), are different. But the modal logician maintains the distinction: the first we shall call a *de dicto* reading of (6) and the second a *de re* reading. The difference becomes important and apparent when we try make the substitution from (5). If we read (6) as (6a), we have analogously,

⁴For instance, see Smullyan 1971.

⁵One distinction that should be mentioned lies in the ambiguity in the expression “. . . expresses a (necessary or contingent) truth”. Consider the sentences P and Q , which are respectively, “ $11 \times 12 = 132$ ” and “Paul is wearing a blue sweater”. Let R be the sentence “ P and Q ”. Does R express a necessary truth? It is surely reasonable to say that it does; R expresses a necessary truth, namely P . On the other hand, if we are to take R as expressing at most one truth, namely the conjunction “ P and Q ”, then R may be taken to express either a contingent truth or a falsity, depending on whether Paul is wearing a blue sweater. But none of the sentences under scrutiny in this essay are conjunctions, so this ambiguity should occasion no disparity in interpretation. However, it is sometimes stylistically useful to use such an expression, and I have done so in the hope that there is no added confusion.

(7a) “The number of planets is necessarily greater than 7” is true *if and only if* “the number of planets is greater than 7” is necessarily true.

Hence, a de dicto reading of (7) is clearly false. But if we read (6) as (6b), we have

(7b) “The number of planets is necessarily greater than 7” is true *if and only if* “there is a unique number of planets, which is necessarily greater than 7” is true.

(7b) is now true. The substitution is truth-preserving, and hence Quine’s argument does not disallow the use of modal operators, provided we read (7) as (7b).

But apparently, this distinction is said to commit us to an acceptance of de re properties, which some think lead to essentialism. I shall discuss such matters later, but I will remark that since essentialism is unpalatable to many philosophers, we must feel a little compelled to answer such a worry. Nonetheless I think the distinction is a rather natural one. I will defend the distinction without worrying whether it actually commits us to the disreputable doctrine of essentialism. Suppose there were no such distinction between (7a) and (7b). That is, suppose their right hand sides to be equivalent.

(8) “The number of planets is greater than 7” is necessarily true *if and only if* “there is a unique number of planets, which is necessarily greater than 7” is true.

This claim, if true, must be an instance of some valid schema. Richard Cartwright suggests that it might be something like the following.⁶

⁶Cartwright 1987, p.142. Notation and numbering have been altered to conform with this essay.

(A) If α is any singular term and ϕ any predicate expression, $\ulcorner \alpha \text{ is } \phi \urcorner$ is necessarily true *if and only if* \ulcorner there is a unique α , which is necessarily $\phi \urcorner$ is true.

Cartwright then asks us to consider the sentences S_1 and S_2 .

S_1 : There is a unique proposition at the top of page 210 of *Word and Object*, which is necessarily true.

S_2 : The proposition at the top of page 210 of *Word and Object* is true.

By (A), S_2 is necessarily true if and only if S_1 is true. But, of course, any proposition could have been expressed at the top of page 210 of *Word and Object*, so S_2 is not necessarily true. That is, “The proposition at the top of page 210 of *Word and Object* is true” is not necessarily true. And we know that S_2 is false without having to know what is actually written at the top of page 210. If there is no de re/de dicto distinction, that is if (A) is correct, we know what to say of S_1 by (A)’s biconditional form and the falsity of S_2 : S_1 is false. So (A) gives us a peculiar result. We can know that S_1 is false without having to know what proposition is being expressed at the top of page 210 of *Word and Object*. This is odd; one would think that the truth-conditions of S_1 depended on what proposition is being expressed. This is not the least of the problems. If we know what proposition is being expressed, we will learn that S_1 is certainly true. The proposition expressed at the top of page 210 of *Word and Object* is that “for every positive integer x , the class of positive integers less than or equal to x has x members”, and this is, of course, necessarily true.⁷ So S_1 is true despite S_2 being false. Thus, (A) must be wrong; and there is a much needed distinction between de re and de dicto modal assertions.

⁷Cartwright 1987, p. 142

Some may feel that this way of addressing the worry—by invoking distinctions and citing ambiguities—is more avoiding than answering. It changes the subject. Moreover, it still stands that substitution in the argument (5) to (7) leads to a falsehood; it is of little value to speak of other sentences when we are trying to judge whether this falsifies the law of substitutivity, making the \Box operator illegitimate. Of course, this objection, though natural, fails to see the point of the *de re/de dicto* distinction. What drives the proponent of this distinction is that sentences are often ambiguous, and to speak of the truth value of an ambiguous sentence is to speak without a subject. The opponent is driven by the useful fiction that sentences are unambiguous, and each expresses at most one proposition.

The argument as presented in (5)–(7) can be reformulated in the following way to obviate any ambiguity. We can still arrive at a falsehood when we observe the *de re/de dicto* distinction. Let us read (5) to (7) in a thoroughly *de dicto* fashion.

(5) The number of planets is 9.

(6c) It is necessary that: 9 is greater than 7.

Notice that by substitution we arrive at

(7c) It is necessary that: the number of planets is greater than 7.

Hence, a falsehood. Treating necessity in these arguments thoroughly *de dicto*, we seem bound to arrive at a false conclusion. Substitution leads to a falsehood, and hence it seems that necessity taken as a sentential operator in this context is illegitimate. But first note that by treating (5)–(7) *de re*, the conclusion is acceptable. If the argument is thoroughly *de dicto*, I argue, it is invalid. Thus (5)–(7) poses no direct threat to the modal operator: if it is *de re*, the conclusion is acceptable; if it is *de dicto*, the conclusion does not follow.

The heart of this reply to the de dicto interpretation of the argument begins with Russell. Russell was concerned to solve puzzles of the following sort:

Now George IV wished to know whether Scott was the author of *Waverley*; and in fact Scott *was* the author of *Waverley*. Hence we may substitute *Scott* for *the author of "Waverley"*, and thereby prove that George IV wished to know whether Scott was Scott.⁸

We can see a similarity in the type of sentences between the ones here and the ones with the necessity operator. Here, the sentences are of the form "Someone wished to know so-and-so", where earlier we had "It is necessary that so-and-so". Russell's solution involved a closer look at the sentence "Scott is the author of *Waverley*". It first appears as if this is a statement of identity, but Russell argues that it isn't. The expression "the author of *Waverley*" is not really a name but a definite description. So we should not treat the sentence in question as

$$s = a$$

where s stands for Scott, and a for the author of *Waverley*. Rather, we should understand it as the more complicated

$$(\exists x)(Wx \ \& \ (\forall y)(Wy \rightarrow x = y) \ \& \ x = s).$$

Here, s remains a name-letter for Scott and Wx means x wrote *Waverley*. Without going into all the other reasons and advantages Russell gave in favour of adopting this form of analysis, we should at least note that it does help in the present case. Given that this formalization is not an identity statement but an existential quantification, we have no reason, on formal grounds alone, to think that substitution of "Scott" for

⁸Russell 1905.

“the author of *Waverley*” is legitimate. Hence we never reach the absurd conclusion that George IV wished to know whether Scott was Scott. (Absurd because, as Russell says, George IV is hardly interested in the law of identity.)

Given this framework, it is easy to reconstruct why Quine’s argument is unconvincing to so many. If the argument is taken to be *de dicto*, it is invalid on grounds Russell gave to solve his problem. The premises of Quine’s argument shall be rendered as follows:

$$(5d) \quad \Box(9 > 7)$$

$$(6d) \quad (\exists x)(Nx \ \& \ (\forall y)(Ny \rightarrow x = y) \ \& \ x = 9)$$

where Nx means x numbers the planets. But given this *de dicto* interpretation of (6), we cannot legitimately substitute “9” for “the number of planets” or vice versa. The inference to (7) is blocked. Thus Quine’s puzzle constitutes no direct attack on quantified modal logic nor the use of modal notions in natural language. But this could hardly be all there is to Quine, and we shall soon look at some of his more indirect attacks against quantified modal logic; but first I’d like to make a few roundabout remarks about identity.

2.2 On identity

2.2.1 The law of substitutivity

A serious worry with Kripke’s proof is not so much a worry about his conclusions as it is with one of his premises: the law of substitutivity. Recall Kripke’s “law of substitutivity”:

$$(1) \quad (\forall x)(\forall y)[(x = y) \rightarrow (Fx \rightarrow Fy)]$$

If this law is false, Kripke's proof may no longer stand, at least not with the same footing. One possible way of understanding (1) is as follows:

- (B) For all expressions α and β , if $\ulcorner \alpha = \beta \urcorner$ is true then substitution of β for α is always truth preserving.

It is well known that unrestricted use of the law of substitutivity leads to falsehoods; let me call (B) "the unrestricted law of substitutivity". One might think that the pair of sentences in one of the examples above—where S_1 : "9 is necessarily greater than 7" and S_2 : "The number of planets is necessarily greater than 7"—falsify (B). But they do not, because S_1 and S_2 are ambiguous. However, another example can be readily produced.

(9) Giorgione was so called because of his size.

(10) Giorgione = Barbarelli

An application of (B) yields the following falsehood.

(11) Barbarelli was so called because of his size.

Sentences (9) and (11) are unambiguous, and they falsify (B).⁹ So the unrestricted law of substitutivity cannot serve as a premise in Kripke's argument, but this is insufficient reason to think his argument fails. (1) then calls for a different interpretation.

2.2.2 The principle of identity

When Kripke uses (1) in his argument for the necessity of identity, what he seems to have in mind is the plain fact that if two objects are identical then if one object has

⁹Cartwright 1987, p. 137.

a property, the other must also have that property. That is, (1) should perhaps be interpreted in the following way:

$$(C) (\forall x)(\forall y)[(x = y) \rightarrow (\forall z)(x \text{ has property } z \rightarrow y \text{ has property } z)]$$

Let us call this the principle of identity. The difference, it seems, between the principle of identity and the law of substitutivity is all the difference between the world and discourse about it. In fact, argument (9)–(11) fails to disprove the principle of identity. Showing this hinges on how we answer the question of what property argument (9)–(11) speaks about. More specifically, what property does Giorgione have that Barbarelli does not?¹⁰ Perhaps the property in question is the property of being-so-called-because-of-his-size. Thus, (9) becomes

$$(12) \text{ Giorgione has the property of being so called because of his size.}$$

We must also accept the following.

$$(13) \text{ Giorgione has the property of being called "Barbarelli".}$$

(12) and (13) together give us,

$$(14) \text{ Giorgione has the properties of being called "Barbarelli" and of being so called because of his size.}$$

And hence, by existential generalization, we must have

$$(15) (\exists x)(x \text{ has the properties of being called "Barbarelli" and of being so called because of his size.})$$

¹⁰The arguments presented here are reconstructions of arguments presented in Cartwright 1987, pp. 137–147.

But how are we to understand (15)? It must surely come out false if we read it as saying that ‘there is someone who is called “Barbarelli” and is so called because of his size’, because the expression “so called” here refers back to Barbarelli. And no one is called “Barbarelli” because of his size. To find a counterexample to the principle of identity, we must find two objects that are identical where one has a property the other does not. Giorgione and Barbarelli are not two such objects.

The most questionable part of the argument (12)–(15) is the inference by existential generalization from (14) to (15). There is an implicit inference of existential generalization from (12) to (16).

(16) $(\exists x)(x \text{ has the property of being so called because of his size.})$

One may have doubts about this inference, but this is precisely what the defender of such a property must commit herself to. If the inference is unwarranted, then it is not legitimate to regard the property in question as a property at all. This argument (12)–(15) can then be seen to be a *reductio* of the assumption that it is indeed a property. Perhaps, I have been cheating. Perhaps it is not the property of being so called because of his size that is in question, but rather it is the property of being called “Giorgione” because of his size. But then this surely does not work as a counterexample to the principle of identity. In fact, it becomes an instantiation of the principle, since Barbarelli, too, has the property of being called “Giorgione” because of his size.¹¹

Hence, an important difference between the law of substitutivity and the principle of identity must be recognized. It should be noted at this point that I have not considered if there are any arguments against the principle of identity nor have I

¹¹The expression “so-called” is pronoun-like. Its proper representation would include a variable; so it’s incomplete as a property specification.

considered any arguments for it. This situation may leave us wondering whether it is legitimate at all, but I will proceed on the assumption that it is a true principle. The principle of identity is not a linguistic principle, but a metaphysical one. It makes a certain claim about how the world is; and it is very difficult (perhaps impossible) to imagine it being false. It simply seems obvious, and almost indisputable, that if two objects are identical then any property of one must be a property of the other. I think that the reason for anyone's disbelief in it originate in a confusion of (C), the principle of identity, with (B), the unrestricted law of substitutivity. Any argument directed against the principle of identity can, I think, only falsify some other principle, like (B).

Given the principle of identity reading of (1), Kripke's "law of substitutivity", a question immediately arises: does his proof establish the necessity of identity?

$$(C) (\forall x)(\forall y)[(x = y) \rightarrow (\forall z)(x \text{ has property } z \rightarrow y \text{ has property } z)]$$

From this, we should, by some principle of substitution, arrive at

$$(3) (\forall x)(\forall y)[(x = y) \rightarrow (\Box(x = x) \rightarrow \Box(x = y))].$$

An instance of this schema would look like, for some a and b ,

$$(a = b) \rightarrow ((\Box(a = a) \rightarrow \Box(a = b)))$$

The inference from (C) to (3) depends on there being an appropriate property. But what exactly is the property that both a and b have?

Perhaps it is the property of being necessarily identical to a . This will not suffice. Because we have questioned the clarity of the expression "necessarily greater than 7", it is reasonable to conclude that "necessarily identical to a " suffers from the same ambiguity. Moreover, accepting it as a property seems to carry with it a tacit

acceptance of essentialism. For saying that an object, if it exists, must have such and such a property is dangerously characteristic of an essentialist's beliefs. Perhaps an object called " ϕ " has that property just in case " $\Box(a = \phi)$ " is true. But as a consequence of what we said earlier, not every sentence with a variable describes a property. Recall " ϕ is so called because of his size". It has an occurrence of a variable, but as was shown above, it fails to describe any real property. Perhaps an object ϕ can have this property just in case the object itself satisfies " $\Box(a = \varphi)$ " where φ is the free variable. This, however, raises metaphysical uncertainties about whether it makes any sense for an object to satisfy a modal condition. In particular, can it be said, without metaphysical worry, that an object satisfies the condition: it is necessary that a is identical to it? What would be most desirable is to sidestep these issues and find some other principle that will let Kripke derive his conclusion. Indiscernibility principles may provide just that.

2.3 Indiscernibility principles

In explaining Quine's thoughts, Richard Cartwright provides an argument for the referential opacity of some sentential operators including the modal ones. He begins by expressing serious doubt over the importance of the principle of identity, and claims indiscernibility principles are more fundamental.¹²

If α and α' are distinct variables and ϕ and ϕ' are open sentences alike save that ϕ' has α' free at one or more places at which ϕ has α free, then a universal closure of $\ulcorner (\forall\alpha)(\forall\alpha')(\alpha = \alpha' \rightarrow (\phi \rightarrow \phi')) \urcorner$ is an *indiscernibility*

¹²His earlier paper "Identity and Substitutivity" places importance on the principle of identity but he changes his mind in this paper "Indiscernibility Principles"; see Cartwright 1987, pp. 201–215.

*principle.*¹³

A sentence having this form is an indiscernibility principle only if its variables are objectual and unrestricted in range. Consider the two following indiscernibility principles.

(17) $(\forall x)(\forall y)(x = y \rightarrow (\forall z)(z \text{ is a property of } x \rightarrow z \text{ is a property of } y))$.

(18) $(\forall x)(\forall y)(x = y \rightarrow (\forall z)(z \text{ is a friend of } x \rightarrow z \text{ is a friend of } y))$.

He then tries to argue that these sentences are true in virtue of being indiscernibility principles and that one should not place more emphasis on one over another. He claims that indiscernibility principles are fundamental to objectual quantification, and so also wishes to affirm that all indiscernibility principles are true. As I see it, if all indiscernibility principles are true, then we can sneak in an indiscernibility principle which allows Kripke to draw his conclusion. The principle might be:

(3) $(\forall x)(\forall y)(x = y \rightarrow (\Box(x = x) \rightarrow \Box(x = y)))$.

The difficulty is that not every indiscernibility principle looks true. This, of course, doesn't seem to help Cartwright or Quine's position. Neither does it help Kripke's. And there is more. Unfortunately, the apparently false indiscernibility principles look an awful lot like (3). A false indiscernibility principle may be reasonably extrapolated from the following:

(19) Charlie knows that Cicero denounced Catiline.

(20) Cicero = Tully

By substitution we arrive at the following falsehood.

¹³Again, notation and numbering have been altered to conform with the style of this essay. Cartwright 1987, p. 201.

(21) Charlie knows that Tully denounced Catiline.

False, because Charlie believes “Tully” refers to somebody else.¹⁴ It now seems rather plausible to infer that the following indiscernibility principle has at least one false instance.

(22) $(x)(y)(x = y \rightarrow (\text{Charlie knows that } x \text{ denounced Catiline} \rightarrow \text{Charlie knows that } y \text{ denounced Catiline}))$

This example is certainly no counterexample to our principle of identity because we can ask what property is it that Cicero has but Tully lacks. We will find no suitable answer as we can see using the method above. What makes this indiscernibility principle, (22), look similar to (3) is the common appearance of a that-clause. Given a statement S, sentences of the sort “Charlie knows that S”, “Mary believes that S”, “It is logically true that S”, or “It is necessary that S”, are all, in one respect, very similar. They all have an S prefaced by a sentential operator.

Suppose for now all indiscernibility principles are true, then there are two obvious replies to the apparently false ones. First, one might say that they are only apparently false, but in fact true despite our natural language intuitions to read them otherwise. That is, despite how it sounds to us, both (19) and (21) are true. This is an unattractive position to take, unless we are willing to give up some deep intuitions about the English language. Secondly, one might say that these apparently false indiscernibility principles are not really indiscernibility principles, but confusions disguised as indiscernibility principles. Quine and Cartwright adopt the latter. But if Cartwright and Quine are understood as trying to dispel Kripke’s proof or one of his conclusions, (3), rather than going about this so indirectly, would it not be easier to

¹⁴I have no wish to expound on any theory of belief or meaning, so my exposition here might be a little unsatisfying.

show our indiscernibility principle is false by counterexample? The problem is that we cannot do this so easily. As we have seen by invoking the *de re/de dicto* distinction, we cannot substitute descriptions for names so easily, and substituting names for names in modal contexts is not obviously false. Statements such as “It is necessary that Cicero is Tully” or “It is necessary that Hesperus is Phosphorus” are not unquestionably false. Indeed, Kripke has shown how to understand these statements as true, and that all intuitions to the contrary rest on confusing metaphysical issues with epistemological ones. But if we can get rid of all sentential operators, or find good reason to put away modal operators with the some of the objectionable ones, then our proof cannot rest on any indiscernibility principle.

Logical truth as a sentential operator also seems to create false indiscernibility principles. For instance,

$$(\forall x)(\forall y)[x = y \rightarrow (L(x = x) \rightarrow L(x = y))]$$

certainly doesn't seem true. (Here “LS” means: it is a logical truth that S.) Though it is logically true that Cicero is Cicero, this indiscernibility principle gives us the false claim that it is logically true that Cicero is Tully. But why exactly is this last claim false? Why isn't it logically true that Cicero is Tully? Presumably, it is because physical objects themselves do not stand in logical relation to one another. They stand in certain physical relations; some may be taller than others, or have less mass than others. But saying that an object bears some logical relation to another is rather dubious, except perhaps in those cases of mathematics (if one regards numbers as objects; but in any case, nobody regards numbers as physical objects). The standard scientific picture of the physical world tells us that it is a mess of elementary particles subjected to numerous forces. To say of the physical world that some of its relations exhibit logical form seems rather odd to many of us. It is better to say then that

the logical form of a sentence must take into account the expressions within the sentence, and not simply their meanings, whatever that might be. So the sentences “it is logically true that Cicero is Cicero” and “it is logically true that Cicero is Tully” differ in truth value because the two expressions that follow the two “that”s are syntactically different. The first contains two occurrences of the same expression-type, namely “Cicero”, whereas the second does not. This, it is alleged, makes the first true and the second false.

Quine urges a similar explanation for propositional attitude constructions. It might be argued that we cannot take a Russellian view of propositions concerning (19) and (21). If we did, (19) and (21) might be parsed, respectively, as $\langle \text{Charlie, knows, } \langle \text{Cicero, denounced, Catiline} \rangle \rangle$ and $\langle \text{Charlie, knows, } \langle \text{Tully, denounced, Catiline} \rangle \rangle$. But since Russellian propositions are nothing more than the ordered sets of the objects and the relations they have to each other—where syntactic features of a sentence and its components used to express the proposition are disregarded, and hence also the way the objects are named—(19) and (21) must have the same truth value. This, of course, is counter to natural intuitions. And if we are to respect our intuition, we must then read (19) and (21) differently. Remember the reason our example gives for (19) and (21) differing in truth value was precisely that Charlie is unaware that “Tully” refers to the same person “Cicero” does. (Notice that I cannot simply say that Charlie does not know that Cicero is Tully, for then we have only pushed things a step back. We still need an analysis that will make sense of the truth of “Charlie does not know that Cicero is Tully”.) But this explanation must take into account more than just the objects the names refer to, for it must pay attention to the names themselves—in this case, “Cicero” and “Tully” and perhaps whatever else they may involve. We then might say that “Charlie knows that Cicero denounced

Catiline” means that Charlie assents to the verbal expression “Cicero denounced Catiline”. Thus we can explain (21) analogously: Charlie dissents from the verbal expression “Tully denounced Catiline”. Of course, we need not go so far. All we need is that it sufficiently resembles quotational contexts. And hence we can account for the difference in truth value.

Quantifying into quotational contexts is illegitimate, and, as Quine believed, failure of substitution demonstrates it. Suppose the following were true:

(23) Charlie said “Cicero denounced Catiline”.

It would be obviously false to say that we can deduce

(24) Charlie said “Tully denounced Catiline”.

by a mere substitution of co-referential terms. Our approach here is that since ascriptions of propositional attitudes behave in roughly the same way as expressions lying within quotation marks, we will have a good explanation for the disparity between (19) and (21). But what is still wanting is an explanation of why quantifying into quotational contexts is illegitimate. So, I will try to give Quine’s. First, he thought that failure of substitution entailed a failure of existential generalization. For instance, the following is also not deducible from (23).

(25) There is some x such that Charlie said that “ x denounced Catiline”.

This is false, simply because Charlie said no such thing. (25), in part, claims that Charlie uttered “ x denounced Catiline”. But what he uttered is not accessed by the English quantifier “There is some x such that ...”. If one thinks that there might be a way of understanding this as true, Quine offers other examples to show that quantifying into quotation marks makes little sense. Existential generalization upon the true sentence

“Cicero” contains six letters

gives us the nonsensical

$(\exists x)$ “ x ” contains six letters.¹⁵

A variable is meant to supply an object relative to an assignment.¹⁶ But in this case, having the “ x ” lie within quotation marks prevents it from doing the job of a variable. Moreover, the sentence having the name “Cicero” lying within quotation marks is not a sentence about Cicero at all. It teaches us nothing about the person. In both cases, we are only concerned with the orthographic features of the expression lying within quotation marks. If we only had this example, we might think that quotation marks prevented us from reading anything within it to have any reference whatsoever. However, we do have examples such as (23) where it is silly to say that it makes no reference to Cicero, the person himself. For if I reported (23) to you, we could go on discussing about what else Charlie says of Cicero. The point Quine is trying to make is rather that certain contexts—say a quotational one, an ascription of a propositional attitude, or a modal one—prevents a directly referential use of the expressions occurring within those contexts. We can tell when a context behaves in this way, Quine argues, precisely when substitution of co-referential terms yields a falsehood. This makes some sense. Failure of substitution of co-referential terms indicates that the context the substitution takes place in is sensitive to more than just the object the terms refer to. The thought behind this is that if the two terms are co-referential, then the only thing that could account for a failure of substitution must lie in the context.

¹⁵Quine 1980, p. 146.

¹⁶This is so when we are considering an objectual interpretation of the quantifiers, as opposed to a substitutional interpretation of the quantifiers.

An important remark to make is that this points to Quine's belief that indiscernibility principles differ greatly from (B), the unrestricted law of substitutivity. The latter, as we've seen, has false instances, whereas, for Quine, the first does not because a false indiscernibility principle is incoherent. In quantified statements, variables are meant to be purely referential. As Quine writes, 'the positions of "x" and "y" in [indiscernibility principles] must be referential if "x" and "y" in these positions are to be bound by the initial " $(\forall x)$ " and " $(\forall y)$ ".'¹⁷ False instances of the schema cannot then be legitimately quantified over. (We no longer have to suppose that all indiscernibility principles are true; this is something Quine is going to have to argue for.)

2.3.1 Quine's claim of referential opacity

Quine's example with the number of planets was partly intended to show us that modal operators on sentences are sensitive to more than just the objects the expressions referred to—they were sensitive to other semantic or syntactic features of the expressions. Insofar as I invoked a *de re/de dicto* distinction against his argument, this much remains in agreement. Even though Quine's direct attack (his argument with the number of planets) is defused, if Quine is right about indiscernibility principles and objectual interpretation, we still seem to be in something of a bind. In trying to figure out what Quine is doing, it will do us good to be a little careful in stating his position. Referential opacity is sometimes seen as the definition simpliciter of a failure of substitution. We should keep in mind that "referential opacity" is a technical term, but it carries with it a dangerous air of metaphor. It evokes an image of referential expressions being fogged up like windows that were once transparent. This

¹⁷Notation has been altered to conform with this essay. Quine 1960, pp. 167–8.

image should not immediately arise; and heeding this advice is, I think, sound. But let us give Quine the term anyway. Since I take Quine to be saying that it is contexts which are referentially opaque, let me accordingly call the context of an occurrence of an expression α in a sentence S the result of removing an occurrence of α from S . So the context of “9” in “9 is greater than 7” is “— is greater than 7”. Likewise, the context of “Cicero” in ‘ “Cicero” has six letters’ is ‘ “—” has six letters’.

- (D) Given two co-referring expressions α and β , and a sentence S with an occurrence of α , the context of an occurrence of α in S is said to be *referentially opaque* if and only if given a sentence S' , where S' is the result of placing β in the context of that same occurrence of α , S and S' have a different truth value. The context is said to be *referentially transparent* if there is no such pair α and β .¹⁸

Now that we have defined referential opacity, we can more properly state one of Quine’s seemingly innocuous claims. A referentially opaque context prevents the expressions within that context from being purely referential. Perhaps the following will capture what he means:

- (i) Given two co-referring expressions α and β , and a sentence S with an occurrence of α , and a sentence S' which is the result of placing β in the context of that same occurrence of α , the context is referentially opaque if and only if at least one of the expressions α and β is not purely referential. Otherwise, α and β are purely referential.

¹⁸It should be noted that my remarks here are largely inspired by Kazmi 1987. I say “inspired” because they sufficiently differ in detail that whatever faults in logic they possibly present are mine and not his.

This is quite a strong claim. Not only is it false, I think it misunderstands Quine's claim. Consider the following:

S_1 : "Cicero" has six letters.

S_2 : "Tully" has six letters.

S_1 and S_2 differ in truth value. So (i) commits us to claiming that at least one of the expressions, either "Cicero" or "Tully" is not purely referential. But which one is not purely referential? Notice that it is not the occurrence of the expression that fails to be purely referential but the expression itself. This itself is not odd given the familiar distinction between names and descriptions, but this is surely not what Quine means. If it was, the pair of sentences $\{S_1, S_3\}$ would likewise commit us to claiming that at least one of the expressions, either "Cicero" or "Catiline's denouncer", is not purely referential.

S_3 : "Catiline's denouncer" has six letters.

So given (i), and the ability to construct similar contexts like the one of "Cicero" in S_1 , we are likely to conclude that all referential expressions fail to be purely referential. Surely, Quine does not mean to arrive at this conclusion. So we can revise (i) accordingly:

- (ii) Given two co-referring expressions α and β , a sentence S with an occurrence of α , and a sentence S' which is the result of placing β in the context of that same occurrence of α , the context is referentially opaque if and only if the occurrences of the expressions α and β in that context are not purely referential. Otherwise, the occurrences of the expressions α and β are purely referential.

This characterization will also not do. We are convinced, for the most part, that quotation marks prevent a purely referential occurrence of the expression it holds.

Nonetheless,

S_4 : “Giorgione” names a chess player.

is true, and indeed so is

S_5 : “Barbarelli” names a chess player.

We may think that substitution into quotational contexts for co-referring expressions remains illegitimate, but we should know that we have not yet established that. Besides, the substitution in this case yields no falsehood. Still, we are hesitant to say that the occurrences of “Barbarelli” and “Giorgione” in S_4 and S_5 are purely referential. Moreover, (ii), given the sentences S_6 and S_7 , would also tell us that the respective occurrences of “Giorgione” and “Barbarelli” are purely referential.

S_6 : “Giorgione” contains the fifth English letter.

S_7 : “Barbarelli” contains the fifth English letter.

This result is deeply counterintuitive. Better to change (ii) than to give up our intuitions here. Let us again revise (ii) by removing its biconditional form, but in doing so, we should know that we give up the attempt to define what it is for an occurrence of referential expression to be purely referential. For my purposes, this is okay, because I believe that this case must be argued for rather than just stipulated. As the result of thinking through the consequences of (i) and (ii), we can now better formulate Quine’s claim.

(E) Given two co-referring expressions α and β , a sentence S with an occurrence of α , and a sentence S' which is the result of placing β

in the context of that same occurrence of α , if the context is referentially opaque then the occurrences of the expressions α and β in that context are not purely referential.

This claim reflects his remark that “failure of substitutivity reveals merely that the occurrence to be supplanted is not *purely referential*, that is, that the statement depends not only on the object but on the form of the name.”¹⁹ Verifying the soundness of this claim would require some definition of what it means for an occurrence of an expression to be purely referential.

2.3.2 Purely referential

One may object to my use of quotation marks in the arguments above on the basis that such use is illegitimate. But as I remarked earlier, this legitimacy is precisely what we are inquiring about. So in trying to give a definition of pure reference, we cannot simply exclude quotational contexts because we find it neat and tidy. A systematic rejection of such usage in these arguments cannot be made without already knowing what counts as pure reference, even if we must firmly keep in mind our intuitions to the contrary.

An attempt at a definition should, I think, avoid use of the law of substitutivity. I wish to say that an occurrence of an expression in a sentence counts as purely referential if all it contributes to the sentence is its referent. Taking advantage of the notion of satisfaction and the fact that variables are meant to be purely referential given an objectual interpretation of the quantifiers, we can state something like the following:

¹⁹Emphasis is his; Quine 1980, p. 140.

- (i) Given a sentence S and a referential expression α occurring in S , and a sentence S' which is the result of placing a variable in the context of the occurrence of α in S , the occurrence of α in S is said to be *purely referential* if and only if the reference of α satisfies S' .

We should notice that this definition is a little inadequate. Reflection on the sentence S_1 suggests that the occurrence of “Giorgione” in it is not purely referential.

S_1 : Giorgione is a thin man.

S_1 is false, and more to the point, Giorgione does not satisfy S_2 .

S_2 : x is a thin man.

Despite our characterization in (i), and the presence of S_1 and S_2 , claiming that the occurrence of “Giorgione” in S_1 is not purely referential seems premature. But this difficulty seems to arise because we have not paid due attention to the negative respect of the biconditional form of the definition. We could easily enough revise it.

- (F) Given a sentence S and a referential expression α occurring in S , and a sentence S' which is the result of placing a variable in the context of the occurrence of α in S , the occurrence of α in S is said to be *purely referential* if and only if the reference of α satisfies S' or its negation.²⁰

I think Quine’s remarks on what counts as a purely referential occurrence of an expression support this characterization. This view is perhaps responsible for leading Quine to postulate another criterion for purely referential occurrences in a sentence; namely,

²⁰To be on the safe side, this last “or” should be regarded as an exclusive “or”, lest we find certain expressions that satisfy both some sentence and its negation.

existential generalization on the sentence yields truth. Let us have the following true sentence.

Cicero denounced Catiline.

The principle of existential generalization gives us

$(\exists x)(x \text{ denounced Catiline}),$

or equivalently, “there is something such that it denounced Catiline”. Quine writes,

From (9), existential generalization would lead to:

$(\exists x)(x \text{ was so called because of its size}),$

that is, “Something was so called because of its size”. This is clearly meaningless, there being no longer any suitable antecedent for “so called”.²¹

Let us roughly formulate Quine’s claim as such:

- (G) Given a true sentence S and a referential expression α occurring in S , if upon existential generalization with respect to an occurrence of α in S yields a falsehood, then that occurrence is not purely referential.

Quine often suggests further that if existential generalization with respect to an occurrence of a referential expression in a sentence yields a truth, then that occurrence would be purely referential. If he were to say so much rather than merely suggest it, (G) would then be stated in a much stronger biconditional form. Nonetheless, it remains a criterion. It may seem that if we accept these definitions, (D) and (F), and accept Quine’s claim (E) and (G), then we should accept Quine’s most important claim—namely, that, there is no quantifying into opaque contexts. Or more precisely,

²¹Numbering has been altered to conform with this essay; Quine 1980, p. 145.

- (H) Given a sentence S and a referential expression α , if the context of an occurrence of α in S is referentially opaque, then a variable in that context cannot refer back to an objectual quantifier prior to that context.

Upon reflection of the discussion here, one might wish to conclude (H). But I do not think it immediately follows. Moreover, not only do I think there is no constructing an argument for (H) from what I have said here; I think (H) is false.

2.3.3 Quantification

Mark Richard and Ali Kazmi have both convincingly argued that we can indeed quantify into opaque contexts, and that false indiscernibility principles are compatible with an objectual interpretation of the quantifiers.

The truth of an indiscernibility principle, say “ $(\forall x)(\forall y)(x = y \rightarrow (Fx \rightarrow Fy))$ ”, guarantees that Fx and that Fy determine the same extension; this is easily seen since said indiscernibility principle states that each object that satisfies the open sentence Fx also satisfies Fy . We are used to thinking of open sentences in this way. Richard characterizes the Quinean thought like this, ‘How ... could “A(x)” be true of things of which “A(y)” wasn’t true? Mere relettering of an open sentence couldn’t change what it is true of!’²² To dispel the appeal of this thought, we must determine what it is that makes quantifiers and variables objectual. (It might be tempting to say that a variable is objectual if and only if it is purely referential. But given (F), our earlier characterization of what counts as “purely referential”, variables are not the kinds of things to be purely referential—only occurrences of variables are. I wish to say that a variable is objectual just in case it supplies an object relative to an

²²Richard 1987, p. 562.

assignment; that is, the semantic content of a variable is objectual just in case its semantic content is its referent. By separating what it means for a variable to be objectual and what it means for an occurrence of a variable to be purely referential, we are left with a very large gap. One might think that an objectual interpretation of variables guarantees that every occurrence of a variable is purely referential. This, of course, is false. For instance, variables occurring in quotational contexts are not purely referential.) Richard asks, ‘Why is it a fact about objectual quantification, as opposed to a fact about the first-order languages with which we are familiar, that “ Fx_1 ” and “ Fx_2 ” determine the same condition?’²³ Richard then proceeds to provide a Tarski-style definition of truth for a first-order language in which “ $(\exists x_1)(Fx_1)$ ” and “ $(\exists x_2)(Fx_2)$ ” differ in truth value but where the quantifiers remain objectual. Discussing his arguments go beyond the scope of this essay, but experts seem to agree that Richard has demonstrated that indiscernibility principles are not fundamental to an objectual interpretation of the quantifiers.²⁴

In addition, if a variable in a sentence is accessible to quantification, this does not guarantee that the occurrence of that variable is purely referential. For instance consider the two following sentences: S_1 : “ $x = x$ ”; and S_2 : “ $x = y$ ”. These sentences are true relative to the assignment of Cicero to both x and y . But consider S_3 and S_4 , which are, respectively, “ $L(x = x)$ ” and “ $L(x = y)$ ”, where “ LS ” states that “it is a logical truth that S ”. Now notice that S_3 is true whereas S_4 is false relative to the assignment of Cicero to both x and y . If these four sentences are, instead, open sentences with x and y as free variables, then only S_1 and S_3 are true for all interpretations of the variables. This consideration suggests the following account of the sentential operator “ L ”.

²³Richard 1987, p.562

²⁴See Salmon 1995 and Soames 1995.

Given a sentence S with at least one variable; “ LS ” is true if and only if S is true in all interpretations.²⁵

Logical form takes into account the syntactic and semantic features of expressions, and not simply the referents of such expressions. But this is not to say that it pays no heed to it all; indeed, our characterization of L demands that it does. Quantification into it then should not be considered illegitimate. It now makes perfect sense to say both that (26) is false and that its quantifiers are objectual in range.

$$(26) (\forall x)(\forall y)[x = y \rightarrow (L(x = x) \rightarrow L(x = y))]$$

Hence, not all indiscernibility principles are true. It is now obvious how there can be a pair of open sentences that differ merely by a letter that do determine different conditions of satisfaction. Moreover, now that we have a false indiscernibility principle which is understandable and accepted, it is obvious that an inference by existential generalization cannot always be permitted. To infer “ $(\exists x)(L(y = x))$ ” from “ $L(y = y)$ ” is illegitimate because the former is false; the L -operator undoes whatever assignment is given to x by the existential quantifier. But notice that we can, on the other hand, infer “ $(\exists x)(L(x = x))$ ” from “ $L(y = y)$ ”. Despite the L -operator’s undoing of the interpretation, the former is true.

Richard and Kazmi’s arguments show that (H) is false. And I think it is clear that even upon accepting (D)–(G), (H) does not follow. Quantifying into opaque contexts

²⁵The argument here is in part inspired by Kazmi 1992. Again, details have been sufficiently altered, so all errors are mine. “Interpretation” is a technical term common to logic. An interpretation of a sentence in logic is simply an arrangement of assignments of truth values to all component sentences and an arrangement of assignments to all variables within that sentence. So given a language with a domain having only two elements, a and b , there are only two possible assignments to the expression “ $x = x$ ”, yielding two interpretations: “ $a = a$ ” and “ $b = b$ ”. But there are four possible assignments to the expression “ $x = y$ ”, yielding four different interpretations: “ $a = a$ ”, “ $a = b$ ”, “ $b = b$ ”, and “ $b = a$ ”.

is, at least, sometimes acceptable; hence, the existence of some false indiscernibility principles is compatible with objectual quantification.

2.4 On necessity

What began this discussion on indiscernibility principles was my hope that I could sneak in (3), “ $(\forall x)(\forall y)[(x = y) \rightarrow (\Box(x = x) \rightarrow \Box(x = y))]$ ”, as one of the indiscernibility principles, which were supposed to be always true. And now that it turns out that they are not always true, it seems that I might have to begin all over again. But this is not so. Our doubts about (3) began because they resembled the false indiscernibility principles in structure. Now that I have established that these structures are acceptable, and that an indiscernibility principle is true or false on grounds independent of those structures, we are in the clear. What’s left is simply the question of (3)’s truth.

Presumably, I am to think that (3) is false because S_1 and S_2 differ in truth value, despite “the number of planets = 9” being true.

$$S_1: \Box(9 > 7)$$

$$S_2: \Box(\text{the number of planets} > 7)$$

Maybe I am supposed to conclude that the occurrence of “the number of planets” in S_2 is not purely referential, and so think that the context is opaque. Or maybe I should simply conclude that S_1 and S_2 falsifies (3). But recall my characterization of what counts as purely referential. For this case, (F) would state that since the referent of the expression “the number of planets” is 9, and that does, in fact, satisfy $\Box(x > 7)$. If (E) is true, then the context $\Box(\text{—} > 7)$ is referentially transparent. This suggests that (3) always comes out true. So (3) comes out true without reliance

on any law of substitutivity or principle of identity. Kripke's result, if not his proof, can be maintained: true statements consisting of the identity sign flanked by proper names are necessarily true.

A nagging doubt obviously remains, and someone might protest. "But S_2 is *false*. The context is therefore opaque, and something is wrong is with your definitions, conditions, and claims." And here is my equally obvious reply: S_2 is not false. It is only false on one interpretation. It is true in terms of satisfaction; the referent of "the number of planets" satisfies the open sentence " $\Box(x > 7)$ ". It is false on the reading where the expression, "the number of planets", is thought to do something other than merely give its referent (perhaps it gives an intension). What would make this idea completely acceptable is a convincing semantic argument, and not merely a formal one, for claiming that descriptions behave differently under modal contexts than names do. Fortunately, a convincing semantic argument, is exactly, what Kripke famously provided.

Chapter 3

Reference

By thinking about possible world semantics and its relation to natural language, Kripke came up with a forceful attack against the dominant theory of names and reference at the time. This attack was made in his lectures *Naming and Necessity*. The descriptive picture of names, which Kripke attributes to Frege and Russell, suffered a heavy blow. Though I think Kripke's arguments fail to completely defeat the descriptive picture, he accomplished at least two things. First, he shows that certain accounts of the descriptive picture must be wrong. Secondly, he provides an alternative to the descriptive picture which was just as coherent and systematic. But many people accept the descriptive theory of names only because it is so coherent and systematic. The existence of an alternative just as coherent and systematic shows that there is a lot more to ask from a theory of names than simply coherence and systematicity. Along the way, Kripke advances certain semantic claims about names and descriptions; in particular, names are rigid designators while the descriptions are generally not. These semantic claims, if true, help close the door against modal problems seen in the previous chapter. But Kripke is not free from his own objectors.

The objections have been made not merely on linguistic grounds, but also on metaphysical and epistemological grounds. I will explain Kripke's claims and some of the more popular objections. I hope to show that these objections stem from conflating issues that should be set apart.

3.1 Naive descriptive theory of names

The naive descriptive picture of names holds, among other things, that there are no semantic differences between names and descriptions. Moreover, the meaning of a name can be given entirely by a definite description. For instance, the meaning of "Gödel" might be "the discoverer of the incompleteness theorems". "Gödel" might mean something else in some other language; but it should be kept in mind that under this theory, these two expressions do not have the same meaning *in virtue* of having the same referent, but rather "Gödel" *means* "the discoverer of the incompleteness theorems". So then every use of the expression "Gödel" can be exchanged *salva veritate* with the expression "the discoverer of the incompleteness theorems".¹ When we learn that Gödel was in attendance at one of the important meetings of the Vienna Circle, we also learn that the discoverer of the incompleteness theorems was in attendance. This is not to say that these two people are joined at the hip, but rather that these two people are one. It is important to note that neither Frege nor Russell held the naive descriptive picture.

In any case, the naive descriptive theory is still rather powerful. Recall the story about Hesperus and Phosphorus. Suppose that all there is to the meaning of "Hesperus" is "the first star seen in the evening", and likewise, all there is to the meaning

¹Barring, of course, quotational contexts.

of “Phosphorus” is “the last star seen in the morning”. The theory can then give an explanation of why, upon hearing the ancient astronomers declare that Hesperus is Phosphorus, we learn something apart from them simply saying that Hesperus is Hesperus. As it can easily be seen, their identity cannot be deduced by mere reflection on the meaning of the words “Hesperus” and “Phosphorus”. The astronomers had to at least look up at the sky. And so, by the naive descriptive theory, “Hesperus is Phosphorus” is not to be regarded as an a priori truth. This accords with our intuitions.

But the naive descriptive theory is also insufficient. Consider the following questions: What if Gödel decided early in his youth to be a painter instead of a logician? What if Gödel died before coming up with the incompleteness theorems? The point of these what-if questions is that stating them carries with it no incoherence. If either one of the hypotheses in the what-if questions were true, it would be the case that Gödel was not the discoverer of the incompleteness theorems. That is, S_1 is not necessarily true.

S_1 : Gödel is the discoverer of the incompleteness theorems.

But according to the naive descriptive theory, S_1 means the same thing as S_2 and S_3 .

S_2 : Gödel is Gödel.

S_3 : The discoverer of the incompleteness theorems is the discoverer of the incompleteness theorems.

Both S_2 and S_3 are necessarily true, where S_1 is surely not. One might object that S_3 is not necessarily true—that, the discoverer of the incompleteness theorems might not have been the discoverer of the incompleteness theorems. And here, I agree. (There is an ambiguity in S_3 , and reading it one way results in a contingent truth. In

another, it is a necessary truth.) But this does not save the naive descriptive theory. S_2 remains necessarily true, while S_3 and S_1 do not.

The naive descriptive theory of names must be wrong. But this discussion suggests a semantic difference between names and descriptions.

3.2 Rigid designation

An expression is said to be a rigid designator if and only if it designates the same object in all possible worlds. Or more precisely,

- (A) An expression α that designates an object x is a *rigid designator* if and only if α designates x in all possible worlds where x exists, and α does not designate something other than x in possible worlds x does not exist in.

Kripke's important claim is that proper names are rigid designators, while descriptions are, generally speaking, not. For now, let me disregard the claim about descriptions and promise to return to it later. So, Kripke's claim can be read as such:

- (B) Given a proper name α , if α designates x , then α designates x rigidly.

Given my characterization of rigid designation in (A), there are two ways an expression could fail to be a rigid designator. Suppose that α is some expression and it designates x in the actual world.

- (1) α does not designate x in all possible worlds where x exists.
- (2) There is a world where x does not exist, and α designates something else.

Someone might think that (B) must be false because (1) is easily satisfied if α is a proper name. There is a possible world where Socrates was not called "Socrates". In that world, Socrates exists but "Socrates" does not designate him. This, I think, is all fine and dandy, but this possible world does not satisfy (1). The expression α is meant to be used in our language in our world to describe another world. I agree it makes sense to say

Socrates might not have been called "Socrates".

Or, alternatively,

Socrates is not called "Socrates" in some possible world.

Notice that there are no contradictions in saying such things. Both are true. But notice that the first appearance of "Socrates" in each sentence still refers to Socrates, even in a world where he is called otherwise. When I use the name "Socrates" I mean it in my language, and in counterfactual situations, I am wondering what might happen to him.

In defending (B), I have to show that (1) and (2) do not obtain when the referential expression is a proper name. So given some proper name, say "Aristotle", (1) should then say

S_1 : There is someone who is identical to Aristotle, but this person might exist and not have been identical to Aristotle.

This certainly seems false. So (1) does not seem true when α is a proper name. (2) would say

S_2 : There is someone who is identical to Aristotle, but if this person did not exist, then somebody else would be Aristotle.

This also seems false. So then “Aristotle” is a rigid designator. The thrust of this analysis comes largely from Kripke’s intuitive test for what counts as a rigid designator and what does not.

Although someone other than the U.S. President in 1970 might have been the U.S. President in 1970 (e.g. Humphrey might have won), no one other than Nixon might have been Nixon.²

And by using the same method, we learn that descriptions, in general, are not rigid designators. Let us see if (1) is true if we use the description “the last great philosopher of antiquity”.

S_3 : There is someone who is identical to the last great philosopher of antiquity, but this person might exist and not be the last great philosopher of antiquity.

This sounds true. The person who was in fact the last great philosopher of antiquity might not have been; he could have decided to never go into philosophy or someone else might have come along after him. Moreover, (2) also seems true when we use this description.

S_4 : There is someone who is identical to the last great philosopher of antiquity, but this person might not exist and somebody else would be the last great philosopher of antiquity.

This certainly sounds true. The last great philosopher could have been Plato instead of Aristotle. It seems then that descriptions fail to be rigid designators. Now one might say that I have only shown that one description is nonrigid and only one proper name is rigid. I still have left all the other descriptions and names to consider. But I

²Kripke 1980, p. 48.

do not think that there is anything unique or peculiar to the description or the name I used. Reflection of this kind on ordinary referential descriptions and proper names, I think, would yield the same conclusions. So I do not think that this objection merits much attention. I will however submit that there may be cases of referential descriptions that do rigidly designate their referents.³

3.2.1 De facto and de jure

Consider the expression “the square of three”. Is it a rigid designator? By replacing α with this expression, we arrive at S_1 and S_2 .

S_1 : There is a number that is identical to the square of three, but this number might exist and not be identical to the square of three.

S_2 : There is a number that is identical to the square of three, but this number might not exist and some other number would be the square of three.

Both seem false. The expression “the square of three” is then a rigid designator. But this expression is rigid for reasons different from why “Aristotle” is rigid. Here, “the square of three” rigidly designates its number, because of the necessary nature of arithmetic. The laws of arithmetic guarantee that certain kinds of arithmetical descriptions of numbers rigidly designate them. The laws of nature, on the other hand, do not guarantee that much. The natural world may turn out to be determined forever, but even if it were not, the mathematical world would remain so.

³Certain kinds of names, like “The Holy Roman Empire”, are not really descriptions. If it were it would be a wholly inaccurate description. As it is well known, the Holy Roman Empire is neither holy, Roman, nor an empire. Other descriptions, such as, “the actual president of the U.S. in 2002” could be considered a rigid designator. The word “actual” in the expression may be said to act like an indexical pointing to the actual world.

My remarks here gesture at a platonic view of arithmetic. None of my arguments, however, depend on any platonistic assumption. The platonic view only comes into play to help me explain why the expression “the square of three” seems rigid. An explanation for this phenomenon is desirable but unnecessary for my purposes. In either case, with an explanation or without, I only wish to assert that despite the existence of some rigid descriptions, generally speaking, descriptions are nonrigid. Nonetheless, our descriptions of the world seem to hang on to their objects ever so loosely. Descriptions generally seem to come off their objects when we ask of them “what if”. But descriptions in mathematics do not. The “square of three” is then called *de facto* rigid for these reasons.

Proper names, on the other hand, are not rigid for these reasons. There are no obvious descriptive features in proper names. They seem to designate what they do without any mediation. For this reason, proper names are thought to be *de jure* rigid designators. If they are, then it seems that almost any descriptive account of proper names must be wrong.⁴ Moreover, in an objectual interpretation of quantified modal logic, interpreted variables are rigid designators. Recall that an interpreted variable supplies an object relative to an assignment. It is by stipulation, and not a matter of semantics, that relative to an assignment, a variable supplies the same object in all possible worlds. Since variables do not refer to their objects via mediation or description, variables are also considered *de jure* rigid. Recall the discussion in the last chapter: this is precisely what guarantees the truth of the following indiscernibility principle.

$$(3) (\forall x)(\forall y)[x = y \rightarrow (\Box(x = x) \rightarrow \Box(x = y))]$$

⁴What counts precisely as *de jure* varies, and so accordingly its consequences on descriptive accounts of proper names.

It has been argued that even though proper names are rigid designators and not, *prima facie*, descriptions, it does not follow that they are *de jure* rigid. They may be some sort of disguised definite descriptions which are *de facto* rigid. Or they may have Fregean senses but this still allows them to act rigidly. Whatever reasons there may be for this claim, I do not think it matters in the present discussion. I am not interested in whether there is some descriptive account of proper names which withstands the objections. I am only interested in the claim that they are rigid designators—*de jure* or otherwise.

3.2.2 Obstinate and persistent

Nathan Salmon has introduced a further distinction which should be made note of. My characterization of rigid designation in (A) leaves open the question of whether α designates x in worlds where x does not exist. Salmon calls rigid designators that do not designate their referents in worlds where they do not exist *persistently* rigid designators. On the other hand, rigid designators that continue to designate their referent even in worlds where its referent does not exist are *obstinately* rigid designators.⁵

Notice that a rigid designator might be both persistently rigid and obstinately rigid, if the referent exists in all possible worlds. The difference between persistently rigid and obstinately rigid turns only on the consequent of the conditional “if x does not exist in a world, ...”. So if “ α ” is a persistently rigid designator, then the “...” will say “then α does not designate x .” If, on the other hand, “ α ” is an obstinately rigid designator, then the “...” will say, instead, “then α (still) designates x .” But if the antecedent is false, that is, if α does exist, then the conditional is true in both

⁵Salmon 1981, p. 34.

cases. So for an expression α that refers to an object that exists in every possible world, the antecedent will always be false. Thus, given an object that exists in every possible world, a rigid designator of that object will be both persistent and obstinate. An example might be “the square of three”. If one is a platonist about numbers, its referent, nine, exists in all possible worlds. Kripke has called such a designator *strongly* rigid.⁶ Likewise, “nine” is also a strongly rigid designator.

Reflection on certain sentences involving proper names will show that proper names are, in general, obstinately rigid designators. Kripke provides the following example.

S_1 : Hitler might never have been born.⁷

For this to be true, the sentence “Hitler was never born” is true under evaluation with respect to some possible world. But this sentence is only true of those worlds where Hitler does not exist. Sentences with non-denoting terms could have truth values, but it is better, I think, to say that they are without truth value at all. Hence, “Hitler” still refers to Hitler with respect to worlds where he does not exist; “Hitler” then seems to be an obstinately rigid designator.

Jason Stanley is unimpressed by this argument. He writes,

It relies on the thesis that sentences that contain non-denoting terms receive no truth-value. If one said that sentences containing non-denoting terms were false, then analyzing “Hitler was never born” as the negation of “Hitler was born” in a world in which “Hitler” is non-denoting would yield the correct prediction.⁸

However, I am partial to the thesis that sentences containing non-denoting terms receive no truth value. If “Hitler was never born” were the negation of “Hitler was

⁶Kripke 1980, p. 48.

⁷ibid., p. 78. He does not say this exactly, but the idea remains the same.

⁸Stanley 1997, p. 567.

born”, then “Hitler was never born” would be “It is not the case that Hitler was born”. But if my thesis is wrong, and Stanley is right that sentences with non-denoting terms are false, it seems that both “Hitler was never born” and “Hitler was born” would be false with respect to that possible world. I cannot see how Stanley can decide, without begging the question, which sentence of the two, “Hitler was born” or its negation “Hitler was never born”, is false and which one is true. And unless one were willing to accept the view that the negation of false sentences are also false, it seems that one should reject the thesis that sentences containing non-denoting terms are false. I am then inclined to say that sentences containing non-denoting terms receive no truth value. So I come down on the side of the obstinate rigidity of proper names. Nonetheless, I think that neither my argument nor Stanley’s is decisive on the matter, because it depends on the status of sentences with non-denoting terms—a vexing enough question on its own. What remains in agreement, and so far unchallenged, is that proper names are rigid designators.

3.2.3 Wide scope

Michael Dummett has raised an argument against Kripke that has been rather popular. Kripke’s argument was supposed to show that there is a semantic difference between names and descriptions. This result, Dummett thinks, is premature. The difference in modal contexts can be accounted for by syntactic considerations alone; any semantic difference is illusory. Dummett claims that names always take wide scope in modal contexts, and descriptions do not. That is, given a first-order logic interpretation of a modal sentence, names sit outside the modal context whereas descriptions do not. For instance, consider the following sentence.

S_1 : Aristotle might not have been the last great philosopher of antiquity.

According to Dummett, we should analyze S_1 as S_2 :

S_2 : For some x and if we let Aristotle = x , then $\Diamond\neg(x = \text{the last great philosopher of antiquity.})$

By the rigidity of interpreted variables, this sentence is true. This accords with our modal intuitions, as I have tried to make clear in discussing Kripke. And so, likewise S_3 and S_4 , its analyzed counterpart, are both false.

S_3 : Aristotle might not have been Aristotle.

S_4 : For some x and if we let Aristotle = x , then $\Diamond\neg(x = x)$.

This syntactic difference even guarantees that Hesperus must be Phosphorus. For S_5 , via its analyzed counterpart S_6 , is surely false.

S_5 : Hesperus might not have been Phosphorus.

S_6 : For some x and some y , if we let Hesperus = x and Phosphorus = y , then $\Diamond\neg(x = y)$.

Along with the fact that Hesperus = Phosphorus, the rigidity of the variables guarantees that S_6 is false, and so then S_5 is false as well. I agree with Dummett that in these sentences, the desirable results can be achieved by consideration of scope distinctions. But Dummett further claims that any appearance of rigid designation is mere appearance, and that it can be completely accounted for by adopting the convention that names always take wide scope.⁹

There are, I think, two problems with this approach. The first is that there needs to be some sort of semantic difference between names and descriptions. We cannot simply stipulate of the English language, like we can with logical languages, that its

⁹Dummett 1973, pp. 113–115.

proper names take wide scope. Even if one were to say that rigid designation can be reduced to or explained entirely by making certain scope considerations, one is still left with a second problem. The second problem is that the approach fails to capture all there is to rigid designation. In answer to Dummett, Kripke provides the following *simple* non-modal sentences for consideration:

S_7 : Aristotle is fond of dogs.

S_8 : The last great philosopher of antiquity is fond of dogs.¹⁰

Kripke's view is that the expression "Aristotle" in S_7 is a rigid designator, whereas the expression "the last great philosopher of antiquity" in S_8 is a nonrigid designator. Claiming that rigid designation can be accounted for by making scope distinctions does not explain this fact. The thesis about rigid designation takes a certain view of the truth conditions concerning S_7 and S_8 with respect to counterfactual situations. In considering the truth of S_7 , we need not think of anyone but Aristotle himself. Kripke writes, "no possible situation in which anyone but Aristotle himself was fond of dogs can be relevant."¹¹ The possible worlds where S_7 is true are therefore not the same worlds where S_8 is true. To make this clearer, consider the following:

S_9 : Aristotle is Aristotle.

S_{10} : Aristotle is the last great philosopher of antiquity.

S_9 is necessarily true, while S_{10} is not. Dummett cannot, I think, explain this fact by any hypothesis about scope conventions. Modal operators are not present in either S_9 or S_{10} ; there is no issue of scope.

Dummett's scope conventions might be able to show S_{11} is true.

¹⁰See Kripke 1980, pp. 6–7.

¹¹*ibid.*, p. 14.

S_{11} : Hesperus is necessarily Phosphorus.

But they fail to show that S_{12} is necessarily true.

S_{12} : Hesperus is Phosphorus

One might say that if we put modal operators on simple sentences, then names take wide scope as well. But the rigid designation thesis, and consideration of these simple sentences, shows that taking narrow scope for the names has the same truth conditions concerning counterfactual situations as taking wide scope. This, however, is quite different from saying that names always take wide scope in the English language.¹² The point is rather that even if we take narrow scope concerning names, we still have our desired results.

I have only so far discussed the objections concerning the technical aspects of rigid designation. There are however metaphysical and epistemological objections that I now wish to consider.

3.3 Transworld identity

One accusation made against Kripke is that his claim about rigid designation (and about much else) presupposes questionable metaphysical and epistemological assumptions. Most notably, his objectors charge that rigid designation presupposes the controversial assumption of transworld identification. A rigid designator is said to designate the same object in all possible worlds. One could hardly accept the claim that proper names are rigid designators if one had doubts about the identity of one object in different worlds. The problems about transworld identification can be considered. I think, in respect to the two following questions: i) what does it mean for

¹²Kripke 1980, p. 12n.

an object to be identical across worlds? and ii) how can we tell if we have the same object in different worlds? Framing the discussion in this way belies some of the subtle relations between these two questions. For instance, one might think that if an object cannot be the same object in different worlds, then there is no need to ask how we can pick out the same object in different worlds. Or, one might think that if there was no way to tell if an object was the same object in different worlds, then it would be useless to suppose that there is any meaning in saying that an object could be the same across worlds.

Certain problems or puzzles have been raised against transworld identification which are particularly sensitive to these questions. In trying to answer these problems, some modalists have raised an analogy between identity across worlds with the older problem of identity across time. I think the analogy could be rather misleading. But insofar as I do think that the thesis of rigid designation presupposes transworld identification, I defend Kripke's position by trying to show that these problems arise from confusions about identity and about the nature of possible worlds.

3.3.1 Three problems

First problem. Let me now state a problem with respect to the first question. I think it is fairly reasonable to suggest that for an object to be the same object across worlds, the object must exist in more than one world. This may sound controversial already. Physical objects do not seem to be the kinds of things that exist in more than one world. Their existence might confine them to any one given place at any one given time. Suppose for now that saying that an object exists in more than one world is okay. Problems still arise. Let x be some physical object that exists in the two possible worlds w_1 and w_2 . First, it seems that we have two objects instead of one;

they could be called, respectively, x -in- w_1 and x -in- w_2 . What we wish to maintain is that x -in- $w_1 = x$ -in- w_2 . But now recall the principle of identity: if two objects are identical then any property of one is a property of the other.¹³ It seems that transworld identity violates this principle. In the present case, if x -in- $w_1 = x$ -in- w_2 , then any property of x -in- w_1 should also be a property of x -in- w_2 .

There is a property that one has and the other does not which immediately stands out. The object x -in- w_1 has the property of being an object in w_1 while the object x -in- w_2 does not. One might say in response that this property arises from a circularity, or one might say that the property in question is not really a property. But another example can be readily produced that is not obviously circular. Kripke wishes to say, for instance, that Nixon might not have been the U.S. President in 1970. Suppose then we have two possible worlds: w_1 is a world where Nixon was the U.S. President in 1970; and w_2 is a world where Nixon was not. Then Nixon-in- w_1 has the property of being the U.S. President in 1970, while Nixon-in- w_2 does not. It seems that if we wish to keep the principle of identity, we should conclude that Nixon-in- w_1 and Nixon-in- w_2 are not identical. Thus the possible world framework seems incompatible with a thesis of transworld identification.

Second problem. Let me now state a problem with respect to the second question I stated in the previous section—how can we tell if we have the same object in different worlds? Consider a possible world, other than the actual one, where Richard Nixon exists. The question can now be reframed in the following way: how do we know

¹³Note that I had labelled this principle (C) in the last chapter. This principle might be better called the “indiscernibility of identicals” to contrast with the identity of indiscernibles, which states that if two objects have the same properties, then they are identical. The conjunction of these two principles has sometimes been called “Leibniz’s law”. But it should be noted that the principle of substitutivity and the set of indiscernibility principles have also been called “Leibniz’s law”. Given the discussion in the last chapter, it should be kept in mind that they say different things. Their confusion results from being the case that, for most situations, all of these principles hold true.

which person in this possible world is Richard Nixon? Would it be enough to say that this individual looked like Richard Nixon? Could we say it is the man with a dog named Checkers, who looked like a certain David Frye impression and happened to be the U.S. President in 1970?¹⁴ In a possible world, Nixon might not have any of these properties. Moreover, in a possible world, somebody else might have all these properties. Would that person then be Nixon? It seems that there is no easy way to tell which person in that world is the one we mean to talk about. There are no criteria of identity for objects across worlds.

Third problem. R. M. Chisholm has presented a third puzzle which takes into consideration aspects of both these puzzles. Consider two different objects in w_0 , the actual world, which we will label x and y . Now it is admitted that these two objects have different sets of properties. Though x and y may have some properties in common x will have properties that y does not. Let us call such properties x -not- y properties. Likewise, y will have some properties x lacks; and so we will call these, y -not- x properties. Suppose then in some nearby possible world, say w_1 , that x fails to have all the x -not- y properties it has in w_0 ; in particular, x loses one x -not- y property to y . This property in w_1 is now a y -not- x property. So let me call this procedure a *transferring* of a property over from one object to another as we move from one world to the next. Now when we move from w_1 to w_2 , y loses a y -not- x property it had in w_0 and it is transferred over to x .¹⁵ So by moving from one world to the next, x and

¹⁴This question is inspired by a passage in Kripke. See Kripke 1980, p. 44.

¹⁵There is a technical difficulty which can easily be accommodated by adopting the following convention. Say in a transfer, y gains a property which is contradictory with one of its other properties. For instance, y -in- w_k is green, and x is supposed to transfer the property of being red upon moving to w_{k+1} . Barring plaid patterns and traffic lights, an object cannot generally be both red and green. (A chess board can be black and white, even black and white all over, but not black all over and white all over.) So when a single transfer will result in an object having contradictory properties, we make two transfers—an exchange, if you will. Thus, x becomes green at the same time y becomes red.

y gradually exchange properties. We continue to move from w_2 to w_3 , and so on in a similar fashion until we reach w_n , where upon it turns out that the set of x -not- y properties in w_n is the same set as the set of y -not- x properties in w_0 . And vice versa: the set of y -not- x properties in w_n is the same set as the set of x -not- y properties in w_0 . The result is that all the properties x has in the actual world are all and the only properties that y has in w_n ; and vice versa. If we allow that transworld identity holds for small changes, we can say that x -in- w_0 is identical to x -in- w_1 , and likewise for each further move. So for each move from a world w_k to w_{k+1} identity is preserved between x -in- w_k and x -in- w_{k+1} . If identity is transitive, then x -in- w_0 is identical to x -in- w_n , and the same is true for y . But what the thought experiment suggests is that x -in- w_0 is identical to y -in- w_n , and that y -in- w_0 is identical to x -in- w_n . Again, by the transitivity of identity, we must conclude that x is identical to y . And that is counter to our initial hypothesis.¹⁶

Moreover, if these are the only changes that occur between the worlds, what difference is there between w_0 and w_n ? Could we tell any difference?

3.3.2 Some proposed solutions

As I have framed them, these problems both have a metaphysical and an epistemological aspect: how an object can exist in more than one world; and how we can tell if we have the same object. Accordingly, how we answer these problems depends on large part on our views about possible worlds. Specifically, they depend on what we think possible worlds are and how we know anything about them. But the distinction between metaphysics and epistemology is never as clear cut as we would like it to

¹⁶My presentation of Chisholm's problem follows Loux 1979, p. 38. For the original problem, see Chisholm 1967.

be. So some of the proposed solutions to these problems address all of them at the same time by offering a full blooded view of possible worlds, complete with a systematic account for answering ontological and epistemological concerns. There are then, roughly speaking, three views about possible worlds. First, we have David Lewis's *modal realist* account. Second, we have Saul Kripke's *modest realist* view. And third, we have the *ersatz modal realist* view, whose proponents include Alvin Plantinga; and they might include Robert Stalnaker, Robert Adams, and Peter van Inwagen.¹⁷ I mention this now only because how these problems will be addressed reflects some of these views.

I shall begin with one possible response to these problems. The first and third problem begin by assuming that an object remains the same if it loses one of its properties. Suppose we say that all properties of an object are essential to its being that object; that is, if an object fails to exhibit one of its properties, then it is no longer the same object. We have then stopped the problems at their common root. But the trouble with this response should be obvious. It concedes too much too quickly. Transworld identity is then only possible if an object in one world is identical in all respects to an object in another. Allowing this would defeat the purpose of possible worlds altogether. In fact, it would suggest that anything that is, must be the case. We can then no longer make sense of ordinary claims like "I wish I had brushed my teeth this morning" without thinking that its speaker is engaging in pure fantasy.

The second response will reflect David Lewis's views. It is to basically deny that transworld identity is really identity at all. To start, individuals in any given possible world are *world bound*. When you say that you might have gotten coffee instead of

¹⁷Lumping these people together might ignore some of the differences in their positions. But I am not too concerned with the details of the third view. There is also a fourth view, which deserves mention, called *modal fictionalism*, whose proponents prominently include Gideon Rosen. See Rosen 1990.

tea this morning, under the modal realist view, what you are actually saying is that your counterpart got coffee instead of tea in some nearby possible world in a morning resembling ours. But what exactly is a counterpart? David Lewis writes,

Your counterparts resemble you in content and context in important respects. They resemble you more closely than do the other things in their worlds. But they are not really you. For each of them is in his own world, and only you are here in the actual world. . . . [The counterpart relation] is the resultant of similarities and dissimilarities in a multitude of respects, weighted by the importances of the various respects and by the degrees of the similarities.¹⁸

Counterpart relations between objects are then based on similarity. The result being that the counterpart relation has different logical properties from that of the transworld identity relation. Whereas the transworld identity is an equivalence relation in virtue of being an identity relation, the counterpart relation is not. For instance, there is no transitivity in the counterpart relation. So in Chisholm's example, *x-in- w_1* may be a counterpart of *x-in- w_0* , and *x-in- w_2* may be a counterpart of *x-in- w_1* , but this is no guarantee that *x-in- w_2* is a counterpart of *x-in- w_0* . Since Lewis stipulates this as part of his theory on counterparts, Chisholm's problem has no force against him. And likewise, the first problem also dissipates. We might wish to frame the second problem as a question of how we are to determine which object is the one we are looking for in another possible world, but Lewis's answer would be that it simply needs to resemble it in sufficient respects. Lewis has then given up actual identity. And so speaking ordinarily of the "identity" of objects in counterfactual situations or in different possible worlds, it is to speak metaphorically in place of speaking actually of counterpart relations. This reflects Lewis's ontological view of possible worlds.

¹⁸Lewis 1983, p. 28.

I believe that there are possible worlds other than the one we happen to inhabit. If an argument is wanted, it is this. It is uncontroversially true that things might be otherwise than they are. I believe, and so do you, that things could mean different things in countless ways. But what does this mean? Ordinary language permits the paraphrase: there are many ways things could have been besides the way they actually are. On the face of it, this sentence is an existential quantification. It says that there exist many entities of a certain description, to wit “ways things could have been”. I believe that things could have been different in countless ways; I believe permissible paraphrases of what I believe; taking the paraphrase at its face value, I therefore believe in the existence of entities that might be called “ways things could have been”. I prefer to call them “possible worlds”.¹⁹

Lewis appears to be casually employing Quine’s criterion of ontological commitment to assert his belief in possible worlds. What he might seem to be doing here, but makes clear later that this is not what he is doing, is saying that “possible worlds” are shorthand for “ways things could have been”. Lewis says, “when I profess realism about possible worlds, I mean to be taken literally.”²⁰ They are what they are, and not something else; they cannot be reduced to anything else. Only under modal realism, I think, would the counterpart relation make sense.

I admit that his view solves the transworld identification problems, but I am not sure it saves rigid designation (of course, it was never meant to). I do not wish to carefully argue against Lewis except to say that his view is simply bizarre. No doubt he knows that many people find his view strange. I willingly fall into the category of people he says stare at him incredulously, thinking “you can’t possibly believe that!” I confess that giving the incredulous stare is not much of an argument, but I do wish to articulate one problem I have with Lewis’s theory. Suppose I regret making a certain decision. I do not believe that I can regret making this decision

¹⁹Lewis 1973. p. 84.

²⁰ibid., p. 85.

if I myself could not have done otherwise. In fact, I could not care less what my counterpart does in his world. I am only concerned with me in my world. But in addition to that problem, the epistemological problems of transworld identification comes up most strongly under Lewis's view. If a possible world is an entire world with its own furniture, I do not know how we can go about figuring out what is what in a counterfactual world. It does not seem like there is any fruitful way of telling anything about a counterfactual world, let alone which object is supposed to be my counterpart. I feel with many people feel that these sorts of difficulties are rather compelling reasons to reject Lewis's modal realism. I admit that Lewis might have elegant responses to these problems. But I think that such a view of possible worlds runs contrary to some of our deepest intuitions on how we think about counterfactual situations. An entire system of all possible worlds, where the actual world is only one such possible world, seems more science fiction than common sense.

There is nothing more I am willing to say about this matter, and I hope I can move on to discuss other proposed solutions to the problems.

3.3.3 By analogy

One way of addressing a new philosophical problem is to reduce it to another older problem. If this other problem has yet to have been satisfactorily solved, then we can say that there was nothing intrinsically wrong with the extra concepts and theories that brought forth the new problem. So it has been suggested that these problems of identity across worlds are, in this way, sufficiently analogous to problems of identity across time. Then the problems are said to arise not as a result of positing concepts like possible worlds but as a result of translating into new terms an older problem about the identity of objects.

An analogy of the first problem. Physical objects for the most part persist through time. I put my notebook in my bag this morning and now it is on my desk. Since I do not think any magician came along to make a switch, I think that the notebook I have here now is the same notebook I put in my bag this morning. But this too seems to violate the principle of identity. Let me tag the time I put my notebook in my bag this morning as t_1 , and the time I write this t_2 . Now my book, b has different properties at different times. It seems then that the object b -at- t_1 has different properties than the object b -at- t_2 . For instance, b -at- t_1 has the properties of being closed and being held in my hands, while b -at- t_2 lacks these properties and has, instead, the properties of being open and being on my desk. Since we wish to observe the principle of identity, it seems then that we must conclude that b -at- t_1 is not identical to b -at- t_2 . There is, therefore, no *trans-time* identity. Before I address this puzzle, let me make the reduction to a problem of time stronger by making further analogies with the other two transworld identity problems.

An analogy of the second problem. I am out grocery shopping with my friend Judy, and we are buying apples. I am slow at picking apples while Judy is not. She is quick and puts a number of apples in a bag. I, on the other hand, only manage to put one apple in the bag before it is full. Later on, we are thinking of making apple pie, and I wonder which apple is the one I had picked at the grocery store. But given that some of the apples look identical, there is no way to tell. There were no discerning features I remember of the apple I picked. It seems then that there are no criteria of identity for this apple across time.

Something is surely absurd about what this line of reasoning suggests. The answer is, of course, that we simply do not need any criteria of identity. We know for a fact that the apple I had picked is in this bag. Not knowing which apple it is does not

mean that there is no such apple or that we cannot talk about it. If the analogy holds to the aforementioned second problem, then we can see that it is rather silly.

An analogy of the third problem. Suppose we have two people Oliver and Stanley who meet at a time t_1 and immediately become close friends. After spending some time together and telling vivid stories about their past, they no longer know for sure whose memories belong to whom. In fact, Oliver is convinced that some of the things that happened to Stanley happened to him. Moreover, Stanley becomes convinced of this too; some of the stories that he used to believe of himself, he now believes of Oliver. Eventually their memories are indistinguishable, and they believe the same things. After a while, and a Kafkaesque supposition, each starts to look like the other. So at some time t_k , the two of them are exactly identical in appearance and have exactly the same beliefs and memories. But this is not the end of the transition. The two of them argue about something, and they decide to end their friendship. In so doing, Oliver steals Stanley's name, and Stanley steals Oliver's name. After they go their separate ways, Oliver begins to look like Stanley when he first met him at t_1 , and Stanley begins to look like Oliver when he first met him at t_1 . The memories that Oliver tries to recover as his own, are actually Stanley's; and, of course, the opposite is true for Stanley. Oliver goes to Stanley's family, believing himself to be Stanley, and having everyone believe he is Stanley. Stanley goes to Oliver's family and friends, believing himself to be Oliver and having everyone believe he is Oliver. The transition is complete at time t_n . Suppose further that all of this happened in a matter of weeks so we can see no aging. Now we are supposed to allow that Stanley-at- t_1 is identical to Stanley-at- t_n and Oliver-at- t_1 is identical to Oliver-at- t_n , for a transition each day counts as sustaining identity. But what this thought experiment suggests is that Stanley-at- t_n is identical to Oliver-at- t_1 and that Oliver-at- t_n is identical to Stanley-

at- t_1 . By our empirical observations, they would be considered identical. Indeed, can we tell who is really Stanley and who is really Oliver?²¹ So if identity is a transitive relation, we are then supposed to conclude that when they first met, Stanley and Oliver were the same person.

If the analogies hold, and we find them compelling, then we should find that the problem of transworld identity is not a problem for modalists but for anyone who wishes to maintain identity. Suppose for now that the analogies hold. We could either leave the problems of trans-time identity for others to solve or we could say something about them. It is better, I think, to say something about them rather than to leave them to the non-modalist. There may be attractive proposals for identity across time which cannot be carried over for identity across worlds.

3.3.4 Some remarks

What I hope to show is that these puzzles set out in Section 3.3.1 cannot cast any serious doubt onto our notions of time or worlds, but rather these puzzles raise fundamental questions about how we think about objects simpliciter.

Remarks on the first problem. No one, I think, truly believes that objects do not persist through time. Despite what some say in their more philosophical moods, most people believe that they existed yesterday. I myself believe that I existed ten years ago. I was different then; I had the shine of youth. Maybe I still have that—I don't know—but I would like to think that I've matured a little since then. In any case I'm taller now than I was then. But my height is not so essential a part of me that without it I would cease to be me. *I*, and not somebody else, was shorter than

²¹I think the analogy I give here is a little misleading. Oliver and Stanley's genetic traits in my example are not switched. A more accurate analogy can be obtained by simply switching times for worlds in Chisholm's example.

I am now. There is nothing strange about that. A counterpart theorist, of Lewis's strain, might disagree. Since they hold that objects are not, in actuality, identical across worlds, they might hold that objects do not actually persist through time. (I say "might" because they might not buy the analogy of "across time" to "across worlds".) "Persistence" and "identity" might be better understood, and explained away, as *spatio-temporal continuity* and *similarity* respectively. Whatever merits this approach may have, I think it gives up too much. For it seems to me that according to this theory, when I say S_1 , I should be understood to mean S_2 .

S_1 : I was shorter in 1992.

S_2 : Some person who is spatio-temporally continuous with me now is, in 1992, shorter than I am now.

There is, I think, something awfully strange about this approach. And indeed, I have already considered, and refrained from accepting, its analogous solution with respect to worlds.

Another approach, which is less bizarre, is to suggest that properties be *time-indexed*. So the property in question when I say that I was shorter ten years ago might be the property of being-short-in-1992.²² So, without any contradiction, I can have both the property of being-short-in-1992 and the property of being-tall-in-2002. Thinking of properties in such a way avoids violating the principle of identity. And so with possible worlds, we might say that properties are *world-indexed* as well.²³ In some possible world, John F. Kennedy might not have been shot. In the actual world, he was. So John F. Kennedy has the property of being-shot-in-the-actual-

²²This is just shorthand. It is better to say that there is no property of being-short-in-1992 over and above having the property of being 4 feet tall in 1992 and having less height than the average 14 year-old male.

²³See Brennan 1988, p. 132.

world and he has the property of being-shot-in-a-possible-world-*w*. There is nothing contradictory in doing that. Again, thinking of properties in such a way avoids violating the principle of identity.

I think that the analogy for the first puzzle is a good one. Whatever approach one might have to the analogy of the first puzzle, I think, can be translated to the first puzzle.

Remarks on the second problem. If any of the problems could have any direct impact on the thesis of rigid designation, it is the second problem. It is a direct attack on how we can possibly designate something in another world. But the attack is quite weak. I think that its analogy in Section 3.3.3 is even better than the first one, because we know exactly what to say about it. As the analogy shows, we simply do not need any criterion of identification which helps us establish whether we have the same object or not in another world. When I ask of *it* whether *it* could have been different in certain ways, I am still referring to *it* and not something else. As Kripke says,

Don't ask: how can I identify this table in another possible world, except by its properties? I have the table in my hands, I can point to it, and when I ask whether *it* might have been in another room, I am talking, by definition, about *it*.²⁴

Similarly, but more obviously, of the second analogy, one should not ask how I am going to identify the apple after I shake up the bag, except by its properties. I have the apple now in my hands, I can point to it and I can wonder what will happen to it in the future. But I am talking about it. If, in the future, the apple still exists, I can still talk about it even if I cannot pick it out from a line-up. There is, I think, no difficulty in talking about an object in a possible world, even if I cannot single it out

²⁴Kripke 1980, pp. 52–53.

when I look into that world through a telescope. All I do when I refer to an object in a possible world is ask what might happen to it.

Remarks on the third problem. Chisholm's example is trickier. If the example is to be taken as an argument, it can be construed as a reductio on the assumption that if an object loses or gains a property, the object remains the same object. Generally speaking, we believe that if Ralph, a dog, loses its tail, it is still the same dog. But at some point in time Ralph had a tail, and at some point in time Ralph did not. We could easily time-index these properties; Ralph has the property of being-tailed-at- t_1 and the property of being-tailless-at- t_2 . But Chisholm's example shows that if we keep on removing properties slowly, we eventually lose the object and not just some property. So I am not sure how to attribute a time or world indexed property to an object that does not seem to be there.

Chisholm's argument is intended to be a reductio, but instead of rejecting the assumption that an object remains the same if it gains a property or loses a property, we could reject the assumption that an object is merely the unqualified sum of its properties. He seems to think that we could only reject this latter assumption if we believe that objects have *essential properties*.²⁵ This provides a way out. An object is then the sum of its essential properties and its non-essential properties. If an object loses one of its non-essential properties, it remains the same object. If an object instead loses one of its essential properties, we lose the object with it. This clearly prevents Chisholm's example from obtaining. The difficulty with this approach, as Chisholm points out, is that we cannot tell which properties are essential. At some point on our jaunt from w_0 to w_n , we lose x with one of the properties that is transferred. Chisholm asks, but when exactly? and which property? He then

²⁵Chisholm 1967, p. 5.

continues to deride the notion of an essential property. If there are essential properties to an object, Chisholm says, there is no way of finding out which ones they are.

There are, I think, at least two problems with Chisholm's attempt at closing this gap on his *reductio* argument. First, if there are essential properties, I do not think we are without any way of telling which ones they are. There is a difference between not being able to tell which object is the one we want when we look at the objects in another world by inspecting them empirically and not being able to tell which object is the one we want from our standpoint within the actual world. This difference was shown in regard to the second problem. Once we make this distinction, it becomes clear that we only need to worry about the latter. Consider the following question taken from Kripke: could we imagine that the Queen of England, the woman herself, was born of different parents from the ones from whom she actually came? We could discover that the person who we think of as the Queen of England was not actually born of the parents we usually attribute to her. But that is very different from wondering whether this very woman could have come from anyone but her actual parents. It seems to me, and to Kripke, that anyone who would come from different parents than the ones the Queen actually came from, would not be that very woman. Kripke asks rhetorically, "How could a person originating from different sperm and egg, be this *very* woman?"²⁶ So I think that if we remove the property of a person's origin, we have left the person behind. This, I admit, is nothing more than a gesture at an intuition. Anyone who wonders what the reasons are for believing in the necessity of origin should perhaps ask someone else. All I want to show, and need to show, is that we are not completely devoid of any ability whatsoever to tell whether a property is essential.

²⁶Kripke 1980, p. 113.

The second problem with Chisholm's discussion is that he thinks that if we reject the idea that an object is merely the sum of its properties, we are committed to the idea that an object is the set of its essential properties. It seems to me that both ideas are wrong. This is not to say, however, that the object sits somewhere behind its properties. This seems wrong to me too. I confess I do not have any idea of what an object really is, so that is all the second problem amounts to.

The puzzles of transworld identity and transworld identification, as I have hoped to have shown, pose no real problem to the thesis of rigid designation.

3.4 Possible worlds

Discussion here and the positions I take suggest a certain picture of possible worlds, which I advocate and defend. But advocating and defending a view falls short of making an argument for it. So in outlining the view, all I hope is that it can be seen to be suggested by the discussion above. I also hope it goes the other way—that it can inform and make clearer some of the claims I made. Before I try to explain what I think possible worlds are, let me say what I think they are not. I do not think they are distant planets or that they exist in different dimensions. If they were, I think they would be too far away for us to know anything about. So I reject David Lewis's view. I think it is the picture of possible worlds as whole entities completely qualitatively described, and separate from the actual world, which invites the problems of transworld identification. This is not to say that possible worlds are merely formal devices either. Kripke suggests that possible worlds might be better called "possible states (or histories) of the world" or "counterfactual situations".²⁷

²⁷Kripke 1980, p. 15.

Kripke is, in some sense, a realist about possible worlds. He simply does not think they are what Lewis thinks they are. Kripke buys then notion that possible worlds are just ways things could have been as well, but they are not, in kind, like Lewis's ways-things-could-have-been.

Consider Kripke's example with a pair of dice.

For each die, there are six possible results. Hence there are thirty-six possible states of the pair of dice, as far as the numbers shown face-up are concerned, though only one of these states corresponds to the way the dice actually will come out.²⁸

Suppose, for instance, that on a roll, the first die turns up a 6 and the second one a 5. So thirty-five possible states are not actual. But notice that we do not need to ask what kinds of things these possible states are. Ignoring what else might occur in one of these non-actual possible states, these possible states are possible worlds. The actual state might be characterized by the ordered pair $\langle 6, 5 \rangle$. One of the unactualized states will then be $\langle 5, 6 \rangle$. We do not think of this unactualized state as anything more than an abstract object. We certainly do not think of $\langle 5, 6 \rangle$ as existing in another dimension which we need to look at. The only empirically manifest qualities of the thirty-six possible states of the dice are the sums of the pairs. A question then arises: do we then need criteria for trans-state identity that could help us distinguish between $\langle 6, 5 \rangle$ and $\langle 5, 6 \rangle$? The answer must follow that it certainly does not; this is something we simply stipulate. This, I think, shows that there is no real difficulty in stipulating that we are talking about x and nothing else when we talk of it counterfactually. So it remains that even in the absence of any distinct qualities for us to distinguish one state from the other, we know that the two states are not the same. Possible worlds are then nothing more than these states writ

²⁸Kripke 1980, p. 16.

large. I will grant that we might have to take care in elaborating possible worlds from this point by considering what is metaphysically possible and what is not, for they are total ways things could have been. But nonetheless, this metaphysical account of possible worlds can give no rise to any problem of transworld identity or identification against rigid designation. Indeed, I think the theses of rigid designation and this view of possible worlds support one another.

Chapter 4

Conclusion

Despite what Kripke makes clear in *Naming and Necessity*, some philosophers, I maintain, remain puzzled because of a certain difficulty with evaluating statements with respect to another possible world. Some, for instance, have concocted clever arguments to show that identity statements between proper names, if true, are contingent. Let me consider one such case.

4.1 Wreen's argument

In a paper entitled "Proper Names and the Necessity of Identity Statements", Michael Wreen presents an argument, contrary to that of Kripke, for the contingent truth of the sentence "Hesperus is Phosphorus". First, he invites us to consider the following.

- (1) "Hesperus" designates Hesperus (namely, Venus.)
- (2) "Phosphorus" designates Phosphorus (namely, Venus.)
- (3) "Hesperus" and "Phosphorus" both designate the same object (namely, Venus.)

And here, Wreen introduces, what he calls, the *disquotation principle*:

- (DP) If two singular terms, “A” and “B”, designate the same object, then
 A is B.

So we can conclude, from (3) and (DP), that

- (4) Hesperus is Phosphorus.

There is nothing seemingly controversial, as of yet, in the argument, but there are a few more steps before we prove the contingency of (4). Wreen states that

- (5) (1) is contingent,
 (6) (2) is contingent, and
 (7) (3) is contingent.

All (7) states is that it is contingent that both “Hesperus” and “Phosphorus” designate the same object. Wreen asserts that if we restrict our possible worlds, as Kripke does, to those in which Hesperus and Phosphorus exist and the terms “Hesperus” and “Phosphorus” exist, then it must be the case that (4) entails (3). Hence,

- (8) (3) and (4) are logically equivalent.

Wreen asserts the following claim about logically equivalent statements.

- (MS) If two statements are logically equivalent, then they have the same
 modal status.

From (8) and (MS) we can conclude that

- (9) (4) is contingent.

That is, “Hesperus is Phosphorus” is contingent. The rest of his paper concerns, what he sees as, possible objections to his argument. Wreen thinks his argument is valid, so the only challenges he entertains are those that question premises (5) and (6). And so he provides various arguments for claiming that (1) and (2) are contingent, depending on what theory of names one subscribes to.

I believe that anyone who follows Kripke’s tradition will accept that a name might refer to something other than what it does in our language. As I mentioned earlier, people often give their pets names like “Ludwig Wittgenstein”; moreover, I also think it is possible that Ludwig Wittgenstein’s parents might not have named him “Ludwig”. Though some philosophers find it contentious to claim that (1) and (2) are contingent, it is not a claim I wish to dispute, but rather, readily assent to. However, the assertions, which Wreen treats as most benign in his argument, are not nearly as benign as he thinks.

Let us look more carefully at some of Wreen’s arguments. (8) may be spelled out in the following way:

- (10) “Hesperus” and “Phosphorus” designate the same object if and only
if Hesperus is Phosphorus.

Along with (7) and (MS), we can arrive at (9), that “Hesperus is Phosphorus” is contingent.

Kripke, I think, might simply deny (8); indeed it seems to be a consequent of one of his important claims in *Naming and Necessity*.

Suppose we identify Hesperus as a certain star seen in the evening and Phosphorus as a certain star, or a certain heavenly body, seen in the morning; then there may be possible worlds in which two different planets would have been seen in just those positions in the evening and morning. However, at least one of them, and maybe both, would not have been

Hesperus, and then that would not have been a situation in which Hesperus was not Phosphorus. It might have been a situation in which the planet seen in this position in the evening was not the planet seen in this position in the morning; but that is not a situation in which Hesperus was not Phosphorus. It might also, if people gave the names “Hesperus” and “Phosphorus” to these planets, be a situation in which some planet other than Hesperus was called “Hesperus”. But even so, it would not be a situation in which Hesperus itself was not Phosphorus.¹

Kripke explicitly states that it is possible that Hesperus not be named “Hesperus” and Phosphorus not be named “Phosphorus” while not changing the fact that Hesperus is still Phosphorus. So, for Kripke, it is clearly possible that (3) be false and (4) be true. He denies that (3) and (4) have the same modal status, and he may go so far as to deny the validity of claiming (8) altogether. Wreen’s own case rests partly on the following modus ponens claim: If (3) and (4) are logically equivalent, then they have the same modal status – which follows from (8) and (MS). But this modus ponens could hardly convince anyone following Kripke’s tradition. For anyone who is convinced that (3) and (4) do not have the same modal status, as those following Kripke, Wreen’s assertion that (3) and (4) are logically equivalent is unwarranted. Indeed, they might question the logical equivalence on the very grounds that (3) and (4) have different modal status, that is, if they even admit (MS). Wreen should fail to convince anyone who is already persuaded that (3) and (4) do not have the same modal status, for he gives no independent grounds for their logical equivalence, insofar as he gives any grounds at all. But likewise, the Kripkean beliefs alone cannot possibly persuade Wreen, and so I will provide several arguments to show that Wreen’s argument fails, the disquotaton principle misleads, and that (MS) needs severe qualification.

Firstly, let me provide the following analogue. If I replace the object and the singular terms in Wreen’s argument with a different object and different singular

¹Kripke 1980, pp. 108–109.

terms, the argument leads to an absurdity.

- (11) "The square of three" designates the square of three, namely nine.
- (12) "Nine" designates nine.
- (13) "The square of three" and "nine" designate the same object, namely nine.

By the disquotation principle, we have

- (14) The square of three is nine.

For the same reasons Wreen gives above, (11) and (12) are contingent. Our language could be different so as to make the term "the square of three" refer to something else. Surely, we could have used the words, "three" and "nine" to refer to numbers other than the ones to which they actually do refer. It would thus also be contingent that "the square of three" and "nine" designate the same object. So (13) is contingent. Given similar restrictions on possible worlds about numbers, we must have that (14) entails (13). And since we already have it that (13) entails (14),

- (15) (13) and (14) are logically equivalent.

If Wreen is right about logical equivalences preserving modal status, then (14) must also be contingent, which is patently absurd. The square of three must be nine. If there is necessity anywhere, it is in mathematics. Wreen's argument must be fallacious. It is now incumbent on us to find out where his argument has gone wrong.

An instance of the use of the disquotation principle in Wreen's case must be something like the following.

- (16) If two singular terms "Hesperus" and "Phosphorus" designate the same object, then Hesperus is Phosphorus.

Now (16) seems both trivial and obvious, but I will contend otherwise. If (16) is to express a true proposition, an appropriate translation of it should also express a true proposition.

(17) Si deux termes, "Hesperus" et "Phosphorus" designent le même objet, l'Hespère est la Phosphère.²

Here, the sentence is in French; and it is neither trivial nor obvious. Specifying which conditions needed to be satisfied in order to make (17) true is not obvious either. What this suggests, I think, is that (16) needs reformulation. It sometimes proves difficult to express a proposition clearly; in particular, if the proposition is about the terms of the language the proposition takes its expression in. Though any proposition must take form in a language, any expression of it cannot be limited to any one language. So I shall reformulate (16) as follows:

(18) If two terms of English "Hesperus" and "Phosphorus" designate the same object, then Hesperus is Phosphorus.

(18) should, and can only, be so understood as expressing the same proposition that (16) does, if (16) expresses any particular proposition at all. The translation of (18) is, accordingly,

(19) Si deux termes en l'anglais "Hesperus" et "Phosphorus" designent le même objet, l'Hespère est la Phosphère.

(19) is still neither trivial nor obvious. But it is true, and it should be argued for.

²For reasons of simplicity (and because I am not sure how they are actually expressed in French.) I've made up the words "l'Hespère" and "la Phosphère" as translations of the English words "Hesperus" and "Phosphorus". One could just as easily use the expressions "étoile du soir" and "étoile du matin" instead, to have "Si deux termes, "Hesperus" et "Phosphorus" designent le même objet, l'étoile du soir est l'étoile du matin". The argument would still work.

- (20) Les deux termes en l'anglais "Hesperus" et "Phosphorus" designent le même objet.
- (21) "L'Hespère" is an appropriate French translation of the English term "Hesperus".
- (22) "La Phosphère" is an appropriate French translation of the English term "Phosphorus".

Then we may conclude that

- (23) L'Hespère est la Phosphère.

Given that (19) expresses the same proposition as

- (24) (20) \rightarrow (23)

and that (16)—reformulated as (18) and then translated as (19)—expresses the same proposition as

- (25) (3) \rightarrow (4)

we may conclude that (19) is true. As it is, (19) requires the truth of (21) and (22) which are contingent facts of the world. Therefore, it should be admitted that (23) does not follow from (20) by logic alone. And by translating back, we should come to realize that (4) does not follow from (3) by logic alone either. Hence neither (24) nor (25) are instantiations of principles of pure logic. So though it first appears that the disquotation principle is one of logic, we now learn that any correct use of it requires contingent facts about the world and, in particular, the language we are using. The disquotation principle should be seen as a nonlogical metalinguistic principle, where its proper use requires close attention the language it is expressed in and the terms it comprises.

At this point, I should like to make a few remarks about logical equivalence. Wreen equates mutual entailment with logical equivalence. Firstly, it should be noted that anything less than logical equivalence will not suffice; entailment, $\varphi \vdash \psi$, on its own cannot demand ψ to have the same modal status as φ . Consider some sentence φ . By disjunction introduction, we can conclude $\varphi \vee \neg\varphi$, hence $\varphi \vdash \varphi \vee \neg\varphi$. φ may be contingent, but what it entails certainly is not. Since Wreen is trying to prove the contingency of (4), having it simply as the *entailment* of (1), (2), (3), and (DP) will not suffice. Let us grant that Wreen is right about (4) entailing (3), and hence we have a logical equivalence. However, to say that two statements are logically equivalent is ambiguous; it could mean one of the following two definitions.

L_1 : φ and ψ are logically equivalent just in case the biconditional $\ulcorner \varphi \text{ iff } \psi \urcorner$ is true.

L_2 : φ and ψ are logically equivalent just in case the biconditional $\ulcorner \varphi \text{ iff } \psi \urcorner$ is logically true.

We should note that the class of logically equivalent statements expressed in L_1 completely contains the class of logically equivalent statements as expressed in L_2 . As I've argued earlier, the disquotation principle, (8), (25) and other similar statements are not true by logic alone. They are propositions about propositions and terms in the English language. So while (3) and (4) may be logically equivalent, their logical equivalence can only be properly understood under L_1 . But logical equivalences exclusively following definition L_1 , I claim, do not preserve modal status. To go back to my earlier example, I had,

- (13) "The square of three" and "nine" designate the same object, namely nine.

(14) The square of three is nine.

For whatever reasons Wreen gives—ones which I may defend—we can infer that (14) follows from (13) and vice versa, and so have,

(15) (13) and (14) are logically equivalent.

Because we followed the same procedure, (15) must be of the same type of logical equivalence as (8) is of. But in (15), it is obvious that despite the logical equivalence, the two do not have the same modal status as I had argued earlier. Wreen is now left with little reason for affirming that logical equivalences maintain modal status, let alone affirming that *his* logical equivalence does so.

There is, of course, a trivial way of understanding any sentence as false, thus, in some sense, any true sentence is only true contingently. Given any true English sentence, we may construct a language just like English, but where the sentence in question turns out false, if uttered in that language. This is easy enough to do by stipulating that the words in the sentence have different meanings in that language. So “the sky is blue” may be false in a different language where “blue” meant green; or “I am me” may be false if “me” meant you. In fact, English itself may have turned out that way; the relationships between all the definienda and the definitia are contingent. But to claim that a true sentence is contingent on such grounds is to misunderstand the more important aspects of necessity and contingency. It is to commit the fallacy of speaking *in* other possible worlds, rather than *of* them. That, I believe, is ultimately all Wreen trades on.

4.2 Final remarks

Kripke's argument for the necessity of identity-statements involving proper names required two conclusions. First, it should come out true that if two objects are identical then it is necessary that they are identical. Second, proper names are rigid designators. To do this, he had to make use of possible world concepts. And so both of these claims were subjected to numerous objections ranging from charges of misunderstanding the difficulties with modal concepts to charges of misunderstanding basic facts of language. I have considered some of these charges, and tried to show that they are baseless.

As we saw, Quine's modal argument failed to be a direct attack on the meaningfulness of modal logic. The indirect attack failed because it relied ultimately on the direct attack. And so what if there are false indiscernibility principles? The modal indiscernibility principle that Kripke needed, " $(\forall x)(\forall y)(x = y \rightarrow (\Box(x = x) \rightarrow \Box(x = y)))$ ", remained intact. All the fighting really took place somewhere else. We just couldn't see that in the haze. Another objection came from Michael Dummett. He argued that there are no semantic differences between descriptions and names, and that the apparent rigidity of proper names can be explained by adopting scope conventions. But there were two problems with this: there had to be some semantic fact about the English language that forces us to adopt this scope convention; and moreover, not all appearances of the rigidity of proper names could be explained by adopting scope conventions. The existence of simple sentences like "Aristotle was the last great philosopher of antiquity" and consideration of their truth-value with respect to possible worlds showed that Dummett misunderstood Kripke. The problems of transworld identification turned out to be pseudo-problems arising from spurious conceptions of possible worlds and extravagant requirements for identification. Michael Wreen took

an old fallacy of confusing one sense of speaking in a possible world with another sense of speaking in a possible world and hid it in a clever argument to derive the contingency of such identity statements. Though I admit that these are hardly all the objections against Kripke and that I am not entirely sure that Kripke's claim is always right, I do think the burden of proof remains to the objectors. Nonetheless, I think that some of the more popular objections miss their target.

I had rather modest aims for this essay. I had only tried to defend Kripke's claim of the necessity of identity against a few objections: Wreen's argument; the problems of transworld identification; Dummett's argument from wide scope; and Quine's arguments of referential opacity. But in considering these objections, I hope to have shed some light into the areas wherein they reside and show that these monsters in the dark are rather tame and not nearly as dangerous as their noises let on.

Bibliography

- Almog, Joseph, Perry, John, and Wettstein, Howard, editors. *Themes from Kaplan*. Oxford: Oxford University Press, 1989.
- Brennan, Andrew. *Conditions of Identity*. Oxford: Clarendon Press, 1988.
- Cartwright, Richard. *Philosophical Essays*. Cambridge, Massachusetts: MIT Press, 1987.
- Chalmers, David. *The Conscious Mind*. Oxford: Oxford University Press, 1996.
- Chisholm, Roderick M.. "Identity Through Possible Worlds: Some Questions." In *Noûs* 1 1967, Nr. 1, 1–8.
- Dummett, Michael. *Frege: Philosophy of Language*. New York: Harper and Row, 1973.
- Fine, Kit. "The Problem of De Re Modality." In Almog, Joseph, Perry, John, and Wettstein, Howard, editors: *Themes from Kaplan*. Oxford: Oxford University Press, 1989, 197–272.
- Hahn, Lewis Edwin and Schlipp, Paul Arthur, editors. *The Philosophy of W. V. Quine*. La Salle: Open Court, 1986.
- Hale, Bob and Wright, Crispin, editors. *A Companion to the Philosophy of Language*. Oxford: Blackwell Publishers, 1997.
- Hanson, Philip and Hunter, Bruce, editors. *Return of the A Priori*. Calgary: University of Calgary Press, 1992.
- Harnish, Robert M., editor. *Basic Topics in the Philosophy of Language*. Upper Saddle River: Prentice Hall, 1994.
- Jackson, Frank. *From Metaphysics to Ethics*. Oxford: Clarendon Press, 1998.
- Kaplan, David. "Opacity." In Hahn, Lewis Edwin and Schlipp, Paul Arthur, editors: *The Philosophy of W. V. Quine*. La Salle: Open Court, 1986, 229–289.

- Kazmi, Ali Akhtar. "Quantification and Opacity." In *Linguistics and Philosophy* 10 1987, 77–100.
- . "Some Remarks on Indiscernibility." In Hanson, Philip and Hunter, Bruce, editors: *Return of the A Priori*. Calgary: University of Calgary Press, 1992, 167–178.
- Kripke, Saul. "Identity and Necessity." In Munitz, Milton, editor: *Identity and Individuation*. New York University Press, 1971, 135–164.
- . *Naming and Necessity*. Harvard University Press, 1980.
- Leonardi, Paolo, editor. *On Quine: New Essays*. New York: Cambridge University Press, 1995.
- Lewis, David K.. *Counterfactuals*. Cambridge, Massachusetts: Harvard University Press, 1973.
- . "Counterpart Theory and Quantified Modal Logic." In *Philosophical Papers*. Volume 1, Oxford: Oxford University Press, 1983, 26–38.
- Lewy, Casimir. *Meaning and Modality*. Cambridge: Cambridge University Press, 1976.
- Linsky, Leonard, editor. *Reference and Modality*. Oxford University Press, 1971.
- Loux, Michael J.. "Introduction: Modality and Metaphysics." In Loux, Michael, editor: *The Possible and the Actual*. Ithaca: Cornell University Press, 1979, 15–65.
- Plantinga, Alvin. *The Nature of Necessity*. Oxford: Clarendon Press, 1974.
- Putnam, Hilary. "The Meaning of 'Meaning'." In Harnish, Robert M., editor: *Basic Topics in the Philosophy of Language*. Upper Saddle River: Prentice Hall, 1994, 221–274.
- Quine, Willard van Orman. *Word and Object*. Cambridge MA: MIT Press, 1960.
- . "Worlds Away." In *The Journal of Philosophy* 73 1976, Nr. 22, 859–863.
- . *From a Logical Point of View*. Cambridge MA: Harvard University Press, 1980.
- Richard, Mark. "Quantification and Leibniz's Law." In *The Philosophical Review* 96 1987, Nr. 4, 555–578.

- Richard, Mark. *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge: Cambridge University Press, 1990.
- Rosen, Gideon. "Modal Fictionalism." In *Mind* 99 1990, Nr. 395, 327–354.
- Russell, Bertrand. "On Denoting." In *Mind* 14 1905, Nr. 56, 479–493.
- Salmon, Nathan. *Reference and Essence*. Princeton, New Jersey: Princeton University Press, 1981.
- . "Relational Belief." In Leonardi, Paolo, editor: *On Quine: New Essays*. New York: Cambridge University Press, 1995.
- Smullyan, Arthur. "Modality and Description." In Linsky, Leonard, editor: *Reference and Modality*. Oxford University Press, 1971, 35–43.
- Soames, Scott. "Revisionism about Reference: a reply to Smith." In *Synthese* 104 1995, 191–216.
- Stalnaker, Robert C.. "Reference and Necessity." In Hale, Bob and Wright, Crispin, editors: *A Companion to the Philosophy of Language*. Oxford: Blackwell Publishers, 1997, 534–554.
- Stanley, Jason. "Names and Rigid Designation." In Hale, Bob and Wright, Crispin, editors: *A Companion to the Philosophy of Language*. Oxford: Blackwell Publishers, 1997, 555–585.
- Wiggins, David. "Identity-Statements." In Butler, R. J., editor: *Analytical Philosophy: Second Series*. Basil Blackwell, 1968, 47–71.
- Wreen, Michael. "Proper Names and the Necessity of Identity Statements." In *Synthese* 114 February 1998, Nr. 2, 319–335.