71-4694

NIEMI, Gunnar Wayne, 1941-
MODALITY AND SELF REFERENCE.

The University of Michigan, Ph.D., 1970
Philosophy

University Microfilms, Inc., Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

MODALITY AND SELF REFERENCE

by

Gunnar Wayne Niemi

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in
The University of Michigan
1970

Doctoral Committee:

Professor Arthur W. Burks, Chairman
Professor William P. Alston
Associate Professor Lawrence Sklar
Assistant Professor Peter G. Hinman

A B S T R A C T

MODALITY AND SELF REFERENCE

by

Gunnar Wayne Niemi

Chairman: Arthur W. Burks

A well-known version of the antinomy of the liar may be constructed by accepting all instances of the schema

(1) "____" is true if and only if ____

and the empirical identity statement

(2) The first sentence on B is "The first sentence on B is not true"

where "B" abbreviates a description of some particular blackboard.

The antinomy of the liar shows that we can not accept all instances of schema (1) for the use of the predicate "is true".

Certain predicates, such as "is necessary" and "is possible", are called modal predicates because they express what used to be known as "modes" of truth. Theories of modal logic can be interpreted as providing us with schemata for the use of modal predicates. An example of such a schema is

(3) If "____" is necessary, then ____.

In the case of the schema for the use of the predicate "is true" we are enjoined from accepting self referential instances. We would like to know if a similar injunction applies to the schemata for the use of modal predicates.

We may ask our question in the following way. Could there be an arrangement of sentences upon the blackboard B such that someone could derive a contradiction if he accepted the identity statements generated by that arrangement of sentences and the instances of the modal schemata of some theory of modal logic?

We construct a formal theory in which both the self referential instances of modal schemata of Lewis' S5 and the identity statements are expressible. We prove that this theory, which we call SM, is consistent.

The antinomy of the liar can be constructed in formal theories. A theory which has an arithmetic subtheory and a one-place predicate "T(x)" which is such that for any sentence S of the theory we can prove

(4) $T(nr(S)) \leftrightarrow S$

is inconsistent. Here "nr(S)" denotes the numeral of the theory which refers to the number associated with S by a system of Godel numbering. This result is usually refer-

red to as Tarski's Theorem.

Richard Montague has proven a theorem which says that if a formal theory of the sort mentioned above contains a one-place predicate "N(x)" for which we can prove for any sentence S

(5) $N(nr(S)) \longrightarrow S$,

and instances of certain other schemata corresponding to theorems of Lewis' S1, then the theory is inconsistent.

We can obtain our informal version of the antinomy of the liar by "translating" the proof of Tarski's Theorem into ordinary English. It would seem that we ought to obtain a modal antinomy by so translating the proof of Montague's Theorem. In view of the consistency of SM, the proof of Montague's Theorem must employ principles which are not expressible in SM. Whether or not these principles are expressible in ordinary English involves the question of whether or not the syntax of ordinary English is finitely axiomatizable.

We try to account for the fact that the theory SM is consistent while a theory which meets the conditions of Montague's Theorem is inconsistent. One difference between the theories is in the naming conventions which they employ; quotation marks on the one hand and a Godel numbering on the other. In the case of the antinomies

involving truth it makes no difference which naming convention is adopted. In the case of a modal antinomy the difference between using referentially opaque quotation marks and a referentially transparent Godel numbering may well be what determines whether or not such an antinomy can be constructed.

CONTENTS

1. A Description of a Modal Antinomy 1

2. A Justification for our Use of Syntactical Treatments of Modality 7

3. Some Problems in the Formulation and Interpretation of a Syntactical Theory of Modality . 25

4. Montague's Anti-Syntacticalist Thesis 37

5. A Proof of Montague's Theorem 46

6. The "Parity of Naming" Thesis 52

7. The Language L and the Theory SM 59

8. A Specification of a Method for Proving the Consistency of SM 66

9. A Proof that the Cut-Elimination Theorem for Formulas of SM Implies the Consistency of SM . . 77

10. A Proof of the Cut-Elimination Theorem for Formulas of SM 86

11. The I-Extensions of SM 108

12. A Proof that Every I-Extension of SM is Consistent 114

13. Weakening the Parity of Naming Thesis 123

14. Some Problems Involved in Evaluating the Weak Form of the Parity of Naming Thesis 134

Bibliography 144

/1/

A DESCRIPTION OF A MODAL ANTINOMY

No one can spend much time studying what has traditionally passed for logic without encountering the antinomy of the liar in one or another of its several versions. A particularly simple version of this antinomy, or paradox, may be constructed as follows.

Let "B" be an abbreviation for a description which singles out some particular blackboard at some particular time. Imagine Figure 1 to be a faithful representation of the state of the designated blackboard at the designated time.

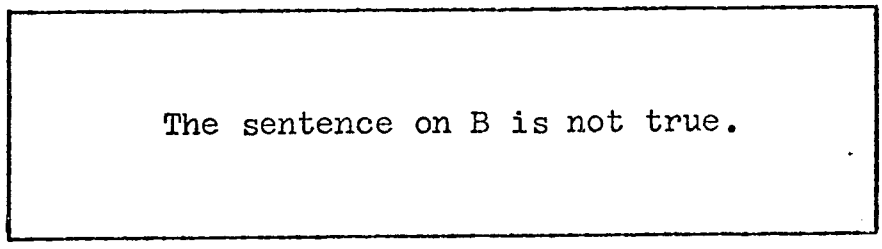


Figure 1

If one accepts the empirical identity statement

(1.1) The sentence on B is "The sentence on B is not true"

and also accepts all instances of the intuitively plausible schema

(1.2) p is true if and only if P,

where we introduce the convention of letting P be a sentence and p be the result of enclosing P in quotation marks, then one has accepted an implicit contradiction.

The contradiction may be made explicit by considering the following instance of (1.2).

(1.3) "The sentence on B is not true" is true if and only if the sentence on B is not true.

Substitution of the identity (1.1) in (1.3) yields

(1.4) The sentence on B is true if and only if the sentence on B is not true.

The antinomy of the liar may be interpreted as telling us something about the use of the predicate "____ is true", namely that we can not accept all instances of (1.2).

If modal operators are replaced by modal predicates, we may interpret the usual formalizations of modal logic as providing schemata for the use of the family of predicates which includes "____ is necessary", "____ is possible", "____ is contingent", and "____ entails ....".

Since schema (1.2) for the use of "____ is true" must be restricted when self reference is afoot, there is a prima facie suspicion that the schemata for using "____

is necessary" and the other modal predicates may require similar restrictions to avoid inconsistency. This suspicion is strengthened when it is recalled that modal predicates are so called because they express what used to be known as "modes" of truth. "____ is necessary", for example, is synonymous with "____ is necessarily true".

So far as is known to the present author, however, there are only three papers, GODEL (1933), KAPLAN and MONTAGUE (1960), and MONTAGUE (1963) which treat of issues connected with the possible need to restrict the use of modal predicates when self reference is involved.

The principal aim of this dissertation is to investigate the consistency of modal schemata which have no restrictions against self referential instances. The emphasis of our study will be on the question of the existence of a general type of modal antinomy, that is, an antinomy analogous to that produced by (1.1). As will be seen later, the three papers mentioned above deal with quite specialized topics.

Let us agree to refer to a person who holds the belief that the modal predicates, like the predicate "____ is true", may lead to a contradiction when self reference is involved as an Antinomist. We will take a Non-Antinomist to be a person who denies this claim. These labels

will be convenient in our later discussion.

The positions set out above have deliberately been left vague. We will attempt to formulate the issue more precisely by giving a description of a state of affairs that might reasonably be construed as establishing the existence of a modal antinomy. The possibility that our specification of what would count as a modal antinomy is not a fair representation of the Antinomist's position will be considered in Section Thirteen.

Consider the following. Suppose that someone agrees to accept all instances of the schemata which correspond to the theorems of some system S of modal logic. Could there then be a set of sentences upon the blackboard B, such that by accepting the identity statements thereby generated plus the instances of the theorems of S, he could derive a contradiction by using valid rules of inference from the first order functional calculus with identity? If there could be such a set of sentences on B then we will say that S has a modal antinomy.

Several different systems of modal logic can be found in the literature. It is possible that some of them have modal antinomies while others do not. Statements about the existence of a modal antinomy, which do not specify a particular system of modal logic, should be understood simply as statements about the existence of a modal anti-

nomy in one or another of the better known systems of modal logic.

It should be emphasized that the contradiction must result from accepting the identity statements generated by the positions of the sentences on B and not from accepting the statements on B themselves. Otherwise a contradiction could be derived by simply letting one of the sentences on B be "Grass is green and grass is not green". The identity statements to be accepted will be of the form "The nth sentence on B is p", where P is in fact the nth sentence on B and p is again the result of enclosing P in quotation marks.

We permit B to contain more than one sentence so that we may allow for the possibility of a modal antinomy involving indirect self reference as well as for the possibility of one involving direct self reference. A sentence may be said to be directly self referential if it contains a referring expression which refers to the sentence itself. A sentence may be said to be indirectly self referential if it refers to other sentences which eventually refer back to the original sentence.

There is a well known antinomy involving the predicate "____ is true" which employs indirect self reference. This antinomy is constructed by arranging for a first sentence to assert the truth of a second sentence

while the second sentence denies the truth of the first.

We conclude this section by presenting, in Figure 2, a mildly complex example of the sort of configuration of chalk upon B which we allow as a candidate for generating a modal antinomy.

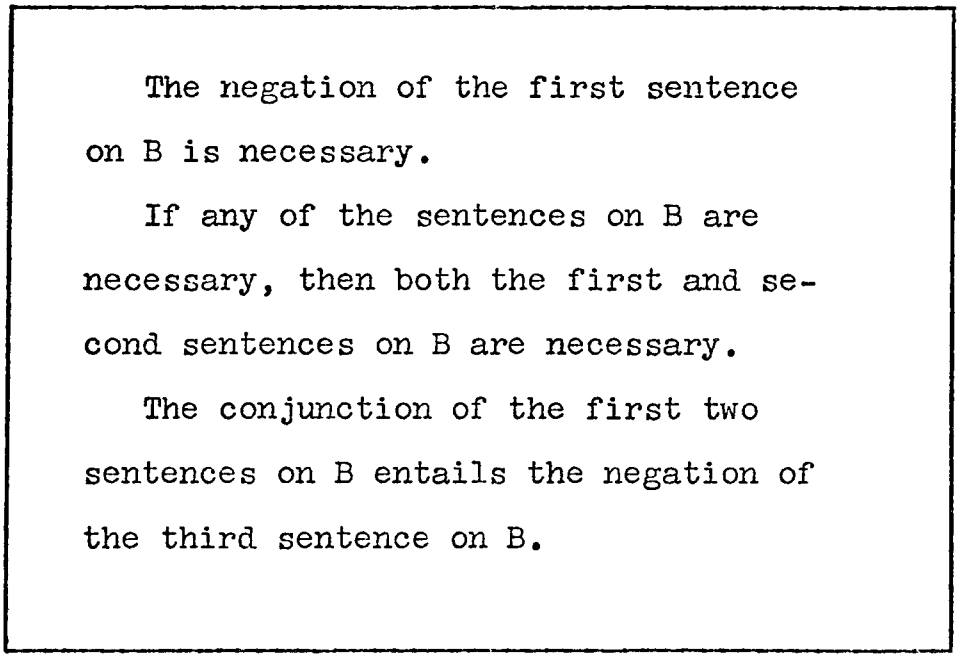


Figure 2

/2/

A JUSTIFICATION FOR OUR USE OF SYNTACTICAL TREATMENTS OF MODALITY

A syntactical theory of modality is a theory which holds that modal sentences, when properly analyzed, will consist of modal predicates, such as "____ is necessary", with names of sentences as grammatical subjects. A non-syntactical theory countenances sentences with modal adverbs, such as "necessarily", inserted into them, or sentences with modal operators, such as "It is necessary that", prefixed to them. We will refer to advocates of these two theories as Syntacticalists and Non-Syntacticalists respectively.

The Syntacticalist will prefer

(2.1) "Nine is greater than seven" is necessary

to either

(2.2) Nine is necessarily greater than seven

or

(2.3) It is necessary that nine is greater than seven.

Syntacticalists point out that their treatment enables one to maintain a strict policy of extensionality. Traditional problems of intensional context, such as the

Morning Star paradox, are resolved by showing that what one might be tempted to take as non-extensional occurrences of a singular term are really no occurrences at all. The occurrence of "Nine" in (2.1) is reduced to an "orthographic accident" like the occurrence of "cat" in "cattle". (2.1) is superior to (2.2) and (2.3) because we are not tempted to replace "Nine" in the former by "The number of planets".

Ohter advantages claimed for the syntactical theory include the avoidance of commitment to an Aristotelian essentialism. QUINE (1953) is a particularly clear account of the philosophical issues involved here.

All of our examples of modal statements in Section One were in conformity with the Syntacticalist's views. We shall continue to confine our attention to the syntactical treatment of modality. Our reasons for limiting our attention, though, do not include a desire for the philosophical benefits which W.V. Quine and others would claim for the syntactical treatment. Being interested in the question of the existence of a modal antinomy we are primarily interested in the technical advantages of using a syntactical treatment of modality.

The reader will do well to recall the simplicity of the derivation of the syntactical version of the antinomy of the liar in Section One, and then to go on to consider

how he would formalize one of the non-syntactical versions of the antinomy, for example, the one which begins

> Suppose a man says that he is lying. If he is lying, then what he says is not the case; but what he says is that he is lying, so it is not the case that he is lying ...

If there is a modal antinomy it is certainly plausible to suppose that self reference is an essential feature of it. To deal with referring expressions we shall want to be able to make use of the traditional logic of identity.

If the modalities are interpreted syntactically, that is as expressing properties of sentences, then the identity calculus may be applied in the usual way. Although some qualifications are needed when distinctions between sentence types and sentence tokens are relevant, the identity of sentences is a reasonably unproblematic notion. To determine whether the same can be said for the identity of entities such as propositions, the reader may look at PRIOR (1963) and evaluate for himself this attempt to determine when propositions are identical.

Our purpose in this section is not to enter into the philosophical dispute between the Syntacticalist and the Non-Syntacticalist. Our purpose here is only to try to provide some reason for supposing that whatever results

are discovered for a syntactical treatment of modality will also hold for a non-syntactical treatment. In particular we want to suggest that the outcome of the dispute between the Antinomist and Non-Antinomist of Section One is not dependent upon whether modal statements are construed syntactically or non-syntactically.

In Section One we constructed a version of the antinomy of the liar which used the syntactical predicate "____ is true". There is a school of thought which maintains that the "mistake" in that argument, that is the feature of the argument actually responsible for the contradiction, is the assumption that truth can be construed as a property of sentences. Since "is necessary" is synonymous with "is necessarily true", it is to be presumed that those who consider it a positive error to treat truth as a property of sentences will also object to the treatment of necessity as a property of sentences.[1]

We wish to argue against the view that an antinomy might result from treating necessity or truth syntactically but that no corresponding difficulty would arise if necessity or truth were treated non-syntactically.

Consider the following "solution" to the version of

1. A comparison of STRAWSON (1948) and STRAWSON (1949), however, may lead one to believe that this is not always the case.

the antinomy of the liar which was presented in Section One.

> But there is a very important difference between our view and that of Tarski. We hold that the adjective 'true' is applicable primarily to propositions, whereas he assumes that it is applicable primarily to sentences. The sentence 'Snow is white' is true, he tells us, if and only if snow is white.[2]
>
> ... And he even goes on to argue that the possibility of constructing the paradox of the Liar within ordinary language shows that for this, as distinct from a formalized language of science, there can be no satisfactory definition of truth. This conclusion is so queer that it should make us suspicious of the assumption from which it is derived; and the source of the trouble seems to be Tarski's unquestioned belief that truth is primarily a property of sentences.[3]
>
> ... It is one of Tarski's main contentions that the paradox of the Liar is due to an inconsistency whereby ordinary language allows for the application of its own words 'true' and 'false' to its own sentences. And he seems to be right in thinking that semantical considerations (as he calls arguments of the kind produced in his paper) require a distinction of language levels. For there is obviously a good sense in which the possibility of constructing the sentence '"Snow

2. KNEALE and KNEALE (1962), p. 588.
3. Ibid, p. 590.

> is white" is true' presupposes the possibility of constructing the sentence 'Snow is white'; and this reflection may be used to explain why the Liar's curious utterance fails to express any proposition at all.[4]

The Kneales' argument does more than just assert that truth is a property of propositions and not of sentences. It also claims to locate the flaw in the reasoning behind the antinomy as the assumption that a certain sentence expresses a proposition when, in fact, it does not. This latter claim is essential. For without it an antinomy could be constructed in a fashion parallel to that of Section One by accepting all instances of the schema

(2.4) p expresses a true proposition if and only if P

and the identity statement

(2.5) The sentence on B is "The sentence on B does not express a true proposition".

The attempt to derive the contradiction will presumably use as an instance of (2.4)

(2.6) "The sentence on B does not express a true proposition" expresses a true proposition if and only if the sentence on B does not express a true proposition.

4. Ibid, p. 590.

The Kneales would object to (2.6) as an instance of (2.4). They would claim that the only legitimate substituends for "P" in that schema are sentences which express propositions and that (2.6) results from (2.4) by the substitution of a sentence which expresses no proposition.

In Section One we interpreted the antinomy of the liar as showing that not all instances of an intuitively plausible schema could be accepted. The Kneales interpret the antinomy as showing that certain sentences do not stand in the expressibility relation to any of the things, propositions, which can be said to be true or false.

Our reply to the Kneales' argument is that we do not see any significant difference between their interpretation and the interpretation of Section One. It seems to us to be just as surprising to discover that a certain sentence does not express a proposition as to discover that one can not accept all instances of an intuitively plausible schema.

The Kneales have not provided any convincing philosophical grounds for distinguishing, in general, between sentences which do and sentences which do not express propositions. It may even be the case that the only grounds that the Kneales have for deciding that a given

sentence does not express a proposition is that the attempt to use that sentence in an intuitively plausible schema leads to a contradiction.

We conclude that the discovery of a modal antinomy by the Syntacticalist will be parallelled by the Non-Syntacticalist's discovery that a sentence, which he had no antecedent grounds for supposing not to have expressed a proposition, did not express a proposition. We also conclude that the latter discovery would be in no less need of an explanation than the former.

Up to this point in our discussion we have assumed that the dispute between the Syntacticalist and the Non-Syntacticalist was a dispute over the sorts of entities of which necessity could be said to be a property. There is, however, another way of interpreting the Non-Syntacticalist's position. The Non-Syntacticalist might be arguing that necessity is not a property of sentences, not because necessity is a property of something different, but because necessity is not a property of anything, either sentences or propositions.

In ANDERSON (1965) the author of that article suggests that necessity is analogous to negation. We shall elaborate upon his suggestion and develop a view which we shall refer to as the operational theory of necessity. Although we deal only with necessity, similar remarks can

be made about the other modalities. —

It seems clear that negation is not a property of anything. The negation of a given proposition is, itself, a proposition. "Snow is not white" expresses the proposition which is the negation of the proposition expressed by "Snow is white". The sentential calculus can be viewed as a set of rules showing, among other things, how propositions and their negations are related.

The operational view would hold, in similar fashion, that "Necessarily, snow is white" expresses a proposition, a modal proposition. Systems of modal logic can be viewed as sets of rules showing, among other things, how the proposition expressed by "Necessarily, snow is white" is related to the propositon expressed by "Snow is white".

There is a good deal to be said for this operational theory. In their informal remarks philosophers like the Kneales, as we saw, talk about necessity as though it was a property. But when they are actually using a system of modal logic they seem to treat necessity operationally. This discrepancy may either be attributed, uncharitably, to a failure to see a distinction or, charitably, to the misleading grammatical form of English idiom. In any case, just as the functional calculus does not allow one to infer

(2.7) (Ex) -x

from

(2.8) -P

so most systems of modal logic would not permit one to infer

(2.9) (Ex) □ x

from

(2.10) □ P. [5]

That the inference of (2.9) from (2.10) is prohibited by the same sorts of considerations that prohibit the inference of (2.7) from (2.8) is an argument in favor of the operational theorist's claim that necessity is analogous to negation.

We will now try to present an argument which an advocate of the operational view might use against the Syntacticalist. The operational theorist will try to argue that the more rational foundations of his own view make

5. Historically, exceptions to this generalization are not lacking. The pioneer work LANGFORD and LEWIS (1932) allows for quantification of propositional variables. Such treatments, however, do not seem to have become very popular. In our opinion they raise foundational questions which they do not satisfactorily answer. See also FEYS (1965), pp. 134-7, for other references to systems of this sort.

it immune from certain difficulties which beset the syntactical theory.

The operational theorist might argue as follows. The Syntacticalist, as well as the Non-Syntacticalist who thinks that modal statements assert that certain propositions have certain properties, is mistaken. Both views, by construing modal statements as being of the subject-predicate form, are susceptible to paradox because they allow the identity calculus to be applied to modal statements. The operational theorist will try to contend that the identity calculus is not applicable to modal statements.

In a modal statement of the form

(2.11) $\Box$ P

no identity substitutions can be made for P since P is a sentence and not a singular term. By appealing to the referential opacity of (2.11), a topic to be discussed in the next section, the operational theorist will also argue that no identity substitutions can be made for occurrences of singular terms within P.

The mistake of the Non-Syntacticalist who does not adhere to the operational view is that he allows (2.11) to be paraphrased as

(2.12) p expresses a necessary proposition.

That is, he treats

(2.13) Necessarily, snow is white

and

(2.14) "Snow is white" expresses a necessary proposition

as equivalent, thus allowing for the substitution of, say, "the sentence on B" for "'Snow is white'" in (2.14). The operational theorist will try to reject the equivalence of (2.13) and (2.14).

Since self reference involves the identity of reference of singular terms, it seems that the only way that a modal antinomy could arise as a result of self reference would be by making identity substitutions in modal statements. This the operational theorist will not allow. Hence, he claims that his view is free of difficulties that the Syntacticalist must face.

We must try to prepare a reply on the Syntacticalist's behalf to the operational theorist's charges. Our line of argument will exploit, still further, the operational theorist's original assumption that necessity is analogous to negation.

Although there is much that is not clear in the shadowy world of propositions, the following seems to be undeniable. If there are such things as propositions, then

we can refer to them. Certainly, in the case of negation, we can refer to the propositon expressed by

(2.15) Snow is not white

by the phrase "the negation of the proposition expressed by 'Snow is white'". There seems to be no reason, in principle, why we shouldn't be able to refer to the proposition expressed by

(2.16) Necessarily, snow is white

by some phrase such as "the necessitation of the proposition expressed by 'Snow is white'".

In the case of negation, whoever accepts

(2.17) Snow is not white $\longrightarrow$ Grass is green

must accept

(2.18) The negation of the proposition expressed by "Snow is white" materially implies the proposition expressed by "Grass is green".

Similarly, whoever accepts

(2.19) Necessarily, snow is white $\longrightarrow$ Snow is white

must accept

(2.20) The necessitation of the proposition expressed by "Snow is white" materially implies the proposi-

tion expressed by "Snow is white".

The identity calculus may be applied to (2.20) even though it can't be applied to (2.19). That is, "'Snow is white'" may be replaced in (2.20) by a phrase such as "the sentence on B".

The operational theorist might try to argue that he is concerned only with the consistency of the modal calculi that he employs and that (2.20) is not a formula of any such calculus. This would be a weak defense. If one is committed to a contradiction it should not matter whether the contradiction is expressible in a particular formal system that one has adopted or only as a metalinguistic statement to which one is committed by having adopted the formal system.

The preceding line of argument is probably derived from an interpretation of TARSKI (1931). There is a way of reading this latter paper so that its author can be taken as saying that ordinary languages are inconsistent, but that's not so bad because they can be abandoned in favor of certain formal languages which may be consistent.[6]

6. See ZIFF (1960), pp. 134-8, for some discussion of this interpretaion of Tarski.

The view that ordinary language may be inconsistent is a silly view. To say that a language is inconsistent might mean that there were proofs in that language both for some statement and its negation and, hence, a proof for any statement whatever. But the standards for what counts as a proof in ordinary language, although not formulated with rigor, certainly rule out arguments which begin "Suppose a man says that he is lying".

We shall conclude this section by showing that the operational theorist is susceptible to a version of the antinomy of the liar which does not use the predicate "____ is true". Insead of "____ is true" this version uses phrases such as "the negation of the proposition expressed by 'Snow is white'". This will create the presumption that if a contradiction can be derived from the use of the predicate "____ is necessary" it can also be derived from the use of phrases such as "the necessitation of the proposition expressed by 'Snow is white'".

We introduce the following notation for our construction of an operational version of the antinomy of the liar. "Tpeb" will abbreviate the phrase "the proposition expressed by". "S" will abbreviate the phrase "the sentence on B". "Implies" will be understood as "materially implies".

We agree to refer to the result of prefixing "It is

not the case that" to a sentence, and changing the initial upper case letter to a lower case letter, as the negation* of the original sentence. We introduce this terminology to satisfy those who may want to insist that only propositions, and not sentences, have negations.

We suppose that our operational theorist agrees to accept all instances of the two schemata

(2.21) If tpeb p implies tpeb q, then if P then Q

and

(2.22) If it is not the case that tpeb p implies tpeb q, then P.

We can arrange for the truth of the identity statement

(2.23) S is "Tpeb S implies tpeb the negation* of S"

from which we can infer

(2.24) The negation* of S is "It is not the case that tpeb S implies tpeb the negation* of S".

Let us assume

(2.25) Tpeb S implies tpeb the negation* of S.

Making the identity substitutions based on (2.23) and (2.24) in (2.25) yields

(2.26) Tpeb "Tpeb S implies tpeb the negation* of S" implies tpeb "It is not the case that tpeb S implies tpeb the negation* of S".

Taking the relevant instance of (2.21) and (2.26) we obtain, by modus ponens,

(2.27) If tpeb S implies tpeb the negation* of S, then it is not the case that tpeb S implies tpeb the negation* of S.

Another use of modus ponens, this time on (2.25) and (2.27), gives us

(2.28) It is not the case that tpeb S implies tpeb the negation* of S.

(2.28) contradicts our assumption, (2.25). Let us reject (2.25) and take (2.28) as an assumption. Making the identity substitutions based on (2.23) and (2.24) in (2.28) yields

(2.29) It is not the case that tpeb "Tpeb S implies tpeb the negation* of S" implies tpeb "It is not the case that tpeb S implies tpeb the negation* of S".

If we apply modus ponens to (2.29) and the relevant instance of (2.22) we obtain

(2.30) Tpeb S implies tpeb the negation* of S

which contradicts the assumption (2.28).

The operational theorist must give up one of the assumptions (2.21)-(2.24). But to decide which one to give up and to explain why it should be given up is just to offer a resolution of the antinomy of the liar.

Thus we conclude by claiming to have made it reasonable to suppose that any problems which a Syntacticalist might have with a modal antinomy are parallelled by analogous difficulties for the Non-Syntacticalist.

/3/

SOME PROBLEMS IN THE FORMULATION AND INTERPRETATION OF A SYNTACTICAL THEORY OF MODALITY

In this section we shall consider certain methodological problems involved in deciding the question of the existence of a modal antinomy for a given system of modal logic.

There are two possibilities. Either a given system of modal logic has a modal antinomy or it does not. In the case where the system does have a modal antinomy there are no problems of method. The question is settled by simply displaying the antinomy. This was done with the antinomy of the liar in Section One. Finding an antinomy is more a matter of ingenuity rather than of systematic technique.

In the case where the system does not have a modal antinomy the question would not be settled simply by one's failure to find an antinomy after some finite amount of time. The question would be settled only by proving that there was no antinomy. A proof of this would require the use of techniques of modern mathematical logic. We shall describe a procedure for carrying out such a proof.

First, a formal language L would be constructed. L

would have to be shown to be an adequate model of the portion of English needed to produce a modal antinomy. That is, it would have to be shown that sentences of the appropriate type, such as self referential modal sentences, were expressible in L. It would also have to be shown that appropriate axioms, such as instances of the modal schemata, and rules of inference, such as identity substitutions, were available for use in L. Finally, the desired theories in L would have to be proven consistent.

Since the formalization of a syntactical theory of modality has never aroused much interest, there is no established opinion as to what the nature of L should be. As the title of this section suggests, the formulation and interpretation of L raises several issues requiring our attention.

A syntactical theory of modality is often referred to as a "metalinguistic" theory of modality. It is also said that the syntactical theory treats the predicate "____ is necessary" as a "metalinguistic" predicate. There is, however, more than one sense in which a predicate may be understood to be a metalinguistic predicate. Our first question, then, is to ask how these claims for a metalinguistic status should affect our construction of the language L.

Consider the following quotations from TARSKI (1931)

and TARSKI (1944).

> ... when we investigate the language of a formalized deductive science, we must always distinguish clearly between the language about which we speak and the language in which we speak, as well as between the science which is the object of our investigation and the science in which the investigation is carried out. The names of the expressions of the first language, and of the relations between them, belong to the second language, called the metalanguage (which may contain the first as a part).[1]
>
> We have implicitly assumed that the language in which the antinomy is constructed contains, in addition to its expressions, also the names of these expressions, as well as semantic terms such as the term "true" referring to sentences of this language; we have also assumed that all sentences which determine the adequate usage of this term can be asserted in the language. A language with these properties will be called "semantically closed."[2]
>
> Since we have agreed not to employ semantically closed languages, we have to use two different languages in discussing the problem of the definition of truth and, more generally, any problems in the field of semantics. The first of these languages is the language which is "talked about" and which is the subject matter of the

1. TARSKI (1931), p. 167.
2. TARSKI (1944), p. 59.

> whole discussion; the definition of truth which we are seeking applies to the sentences of this language. The second is the language in which we "talk about" the first language, and in terms of which we wish, in particular, to construct the definition of truth for the first language. We shall refer to the first language as "the object language," and to the second as "the meta-language."[3]

These quotations describe two different criteria by which an expression may be said to be a metalinguistic expression. The expression may be metalinguistic, first, because it occurs in the studying language rather than the studied language or, second, because it contains the names of linguistic expressions. In Tarski's work these criteria are always found together so no confusion results. There are, however, features of Tarski's approach which make it unsuitable for our purposes.

If the predicate "____ is necessary" must occur in the studying language, then to argue for the absence of a modal antinomy we must rely on the consistency of the studying language.

In TARSKI (1931) that author constructs a definition of the predicate "____ is true" for the elementary calculus of classes. To show the adequacy of his definition

3. TARSKI (1944), p. 60.

he must appeal to the consistency of his studying language which contains a good deal of informal set theory. There is, of course, no doubt but that a portion of set theory which we believe to be consistent could be used to construct his definition of "____ is true". Our point here is that this could be shown only by treating the language in which the predicate "____ is true" occurs as the studied language.

The iteration, or nesting, of predicates also presents a problem on Tarski's approach. If the modal predicates are to occur in the studying language and to take names of the expressions of the studied language as arguments, then we can not formulate sentences with iterated modal predicates.

For these reasons the predicate "____ is necessary" will not be a metalinguistic predicate for us in the sense of being a predicate of th· studying language. It will be a predicate of the studied language. To say this is to say that the studied language will contain a predicate which we intend to interpret as "____ is necessary".

Our modal predicates will, however, be metalinguistic predicates in the sense that they will take as arguments expressions which are to be interpreted as the names of sentences. Since the modal predicates are to appear in the studied language, this language must contain expres-

sions which can be interpreted as names of the sentences of some language or else we would have predicates with no subjects.

As typical examples of modal truths we have the instances of schemata such as

(3.1) p is necessary $\longrightarrow$ P.

If we were interested in constructing a "working" system of modal logic we would want a schema such as (3.1) to be an axiom or theorem of our system. That is, the purpose of our formal system might be to generate all of the schemata whose instances were modal truths. In general, schemata, rather than their instances, would be the theorems of our system. In this respect our formal system would be analogous to the customary formalizations of the sentential calculus.

In these latter formalizations expressions such as

(3.2) Snow is white or it is not the case that snow is white

are not expressions of the formal language. Instead expressions of the formal language are schemata such as

(3.3) P v -P.

Our purposes require that we have more than just schemata such as (3.1) as theorems of our formal system.

It is reasonable to suppose that a modal antinomy, if it existed, would be generated by replacing the sentential letters of modal schemata by sentences which themselves contained modal predicates. In order to represent this procedure in L we would need some technique for forming, within L, expressions which could be interpreted as names of sentences of L when these latter sentences were interpreted as modal sentences in the intended way.

It would seem that we could dispense with sentential letters in L. If the construction of a modal antinomy involved taking an instance of, say, (3.1) in which "P" was replaced by some non-modal sentence, it would be reasonable to suppose that an antinomy could also be constructed by replacing "P" by some arbitrary modal sentence.

Our language L will not be semantically closed according to Tarski's definition of the term since it does not contain the predicate "____ is true". If we describe a language which is semantically closed in Tarski's sense as a language which is semantically closed with respect to truth, we see that L can be described in an analogous fashion as a language which is semantically closed with respect to necessity.

An alternate way of formulating our question about the existence of a modal antinomy would be as the follow-

ing question. Are languages which are semantically closed with respect to necessity, like languages which are semantically closed with respect to truth, inconsistent?

We turn now to the task of providing L with some apparatus for forming names of modal sentences. Up to this point we have followed the practice of forming the name of a sentence by enclosing the sentence in quotation marks. We can simulate this practice in L by allowing L to contain a symbol to be interpreted as a quotation mark. Since we have been using the double quotation mark in our English discussion, we may use the single quotation mark as a symbol of L without fear of confusion.

Thus L might have as a theorem the sentence

(3.4) $N('(Ex)N(x)') \longrightarrow (Ex)N(x)$

which would be interpreted as the English sentence

(3.5) If "There is at least one necessary sentence" is necessary, then there is at least one necessary sentence.

The use of quotation marks to form the names of sentences raises some issues that must be clarified. Thes issues involve what has sometimes been called "the referential opacity" of quotation marks.[4]

Consider the following argument used by Tarski.[5] Suppose that we accept

(3.6) For all p, 'p' is a true sentence if and only if p.

Then he claims that as instances of (3.6) we have both

(3.7) 'p' is a true sentence if and only if it is snowing

and

(3.8) 'p' is a true sentence if and only if it is not snowing.

Tarski goes on to explain that the usual view regarding the expression "'p'" is that it does not contain a quantifiable variable but is a simple expression which denotes the sixteenth letter of the alphabet.

It might be though that if Tarski is right, then we can not express our intentions by saying, for example, that every instance of the schema

(3.9) N('P') $\longrightarrow$ P

4. This term is due to W.V. Quine. A discussion of this type of opacity may be found in the essay entitled "Reference and Modality" in QUINE (1961).

5. TARSKI (1931), pp. 159-60.

is to be an axiom of L. That is, it might be held that the antecedent of (3.9) contained, not a schematic letter for which substitutions might be made, but rather a name for the sixteenth letter of the alphabet.

Upon closer inspection, however, this difficulty can be seen to be illusory. In schemata such as (3.9) we are talking about L, which we are then treating as an uninterpreted language. In particular, we are free to treat the single quotation mark of L as an uninterpreted symbol which is not encumbered by the conventions governing the use of quotation marks in English.

We can acheive generality for the schema (3.9) by referring in our metalanguage, in the sense of our studying language, to all sentences of L which are instances of (3.9). The problems which Tarski encountered were due to his undertaking a task quite different from ours. He was trying to acheive generality within the studied language, which in his case was colloquial English, or Polish, by formulating a sentence of the studied language which would yield all of the desired instances as deductive consequences within the studied language. It was not possible for him to treat quotation marks as uninterpreted symbols.

Although we may treat the single quotation marks as uninterpreted symbols when we are talking about L, we

want their behavior within L to simulate the behavior of quotation marks within English. For example, we do not want

(3.10) $a = b \longrightarrow (N('a = a') \longrightarrow N('a = b'))$,

where "a" and "b" are individual constants of L, to be a theorem of a theory in L. To see that this is undesirable, interpret "a" as "the Morning Star" and "b" as "the Evening Star".

The problem is that (3.10) seems to be an instance of the theorem of the identity calculus,

(3.11) $t = u \longrightarrow (P \longrightarrow P(t/u))$,

where t and u are terms and P(t/u) is the result of replacing a free occurrence of t in P by u in such a way that the new occurrence of u is not bound.

Fortunately this problem can be solved quite easily by altering slightly the existent terminology. All that we have to do is to specify that occurrences of terms within the single quotation marks are bound occurrences. Hence they are not available for identity substitutions.

The preceding stipulation will also prevent quantification over terms which appear within quotation marks. Thus, we will not be able to infer

(3.12) $(Ex)N('N(x)')$

from

(3.13) N('N(a)')

since the occurrence of "a" in (3.13) is not free.

By treating occurrences of terms within quotation marks as bound occurrences we can represent the referentially opaque properties of quotation marks within L.

/4/

MONTAGUE'S ANTI-SYNTACTICALIST THESIS

In this section we shall present a thesis advanced by Richard Montague in MONTAGUE (1963). This thesis, together with the theorem which is used to support it, will provide a point of departure for much of the rest of our discussion.

One of the purposes of MONTAGUE (1963) is stated by that author to be the following.

> On several occasions it has been proposed that modal terms ('necessary,' 'possible,' and the like) be treated as predicates of expressions rather than as sentential operators.
>
> ... The main purpose of the present paper is to consider to what extent within such a treatment the customary laws of modal logic can be maintained.[1]

Syntacticalists have generally tended to favor the weaker systems of modal logic. They would willingly forego some of "the customary laws of modal logic", if by this were meant, for example, "laws of the widely accepted system S5". Many Syntacticalists have doubts about statements such as the following instance of the syntac-

1. MONTAGUE (1963), p. 153.

tical analogue of the characteristic law of Lewis' S4, $\ulcorner \Box P \longrightarrow \Box \Box P \urcorner$.[2]

(4.1) If "All Frenchmen are Europeans" is necessary, then "'All Frenchmen are Europeans' is necessary" is necessary.

The Syntacticalist's objection to the more powerful systems of modal logic has usually been based on a belief that many of the theorems of the syntactical versions of these systems are false, rather than on a fear that they are inconsistent. However, Montague's thesis, which will probably seem extremely harsh to any Syntacticalist, is to be accepted upon pain of inconsistency.

Montague's thesis is expressed in the following quotation.

> Thus if necessity is to be treated syntactically, that is, as a predicate of sentences, as Carnap and Quine have urged, then virtually all of modal logic, even the weak system S1, must be sacrificed.[3]

We will refer to the view expressed in the quotation above as Montague's Anti-Syntacticalist Thesis. To have such a label will be convenient, although possibly mis-

2. See, for example, STRAWSON (1948).
3. MONTAGUE (1963), p. 161.

leading. The thesis, as such, is not yet a criticism of the philosophical foundations of the Syntacticalist's claim that necessity is best treated as a predicate of sentences, but rather a criticism of the formalizations which the Syntacticalist might offer of his treatment of modality. The thesis, if true, would show that the Syntacticalist has overestimated the amount of formal machinery at his disposal. As it stands now, the thesis leaves open the possibility that perhaps only minor revisions are needed to effect a formalization of a reasonably strong syntactical theory of modality.

In KAPLAN and MONTAGUE (1960) the authors of that article present a version of the paradox known variously as the paradox of the Hangman, the Surprise Examination, or the Knower. Their version of this paradox substantiates Montague's Anti-Syntacticalist Thesis, although that is not their only purpose in that article. Of this paradox they say

> Treatments of the paradox have for the most part proceeded by explaining it away, that is, by offering formulations which can be shown not to be paradoxical. We feel ... that the interesting problem in this domain is of a quite different character; it is to discover an exact formulation of the puzzle which is genuinely paradoxical. The Hangman might then take a place beside the Liar and the Richard paradox, and, not unthink-

ably, lead like them to important technical progress.[4]

The condition which leads to the formulation of the genuine paradox is, in effect, the adoption of a syntactical theory of "epistemic" modality as strong as Lewis' S1. Here the authors' conclusion is not that their result shows that an adequate syntactical treatment of knowledge is impossible, but only that certain restrictions are needed.

> There are a number of restrictions which might be imposed on a formalized theory of knowledge in order to avoid the contradiction above. Of these, the simplest intuitively satisfactory course is to distinguish here as in semantics between an object language and a metalanguage, the first of which would be a proper part of the second. In particular, the predicate 'knows' would occur only in the metalanguage, and would significantly apply only to sentences of the object language. According to this proposal, a sentence like 'K knows 'K knows 'Snow is white''' or 'Socrates knows 'there are things which Socrates does not know'' would be construed as meaningless. A less restrictive course would involve a sequence of metalanguages, each containing a distinctive predicate of knowledge, which would meaningfully apply only to sentences of languages earlier in the sequence. A more drastic measure (which

4. KAPLAN and MONTAGUE (1960), p. 79.

> seems to us distinctly unreasonable) is to reject some part of elementary syntax, perhaps by denying the existence of self-referential sentences.[5]

Montague evidently believes that his Anti-Syntacticalist thesis shows that the Syntacticalist has certain difficulties which the Non-Syntacticalist does not share. Immediately after his statement of the Anti-Syntacticalist thesis Montague goes on to say

> This is not to say that the Lewis systems have no natural interpretation. Indeed, if necessity is regarded as a sentential operator, then perfectly natural model-theoretic interpretations may be found ... which satisfy all the Lewis systems S1-S5.
>
> ... It seems at present doubtful that any philosophical interest can be attached to S1-S4. The natural model-theoretic treatment gives a system stronger than all of them, and no satisfactory syntactical treatment can be given for any of them.[6]

Both the above quotation and the statement of the Anti-Syntacticalist thesis suggest that Montague believes that there is something unsatisfactory about the Syntacticalist's position. Presumably Montague believes that it would count as a defect in the Syntacticalist's posi-

5. KAPLAN and MONTAGUE (1960), p. 88.
6. MONTAGUE (1963). p. 161.

tion if he had to accept some restriction like one of those suggested for avoiding the paradox of the Knower. In the case of the restriction prohibiting the iteration of modal predicates this seems quite clear. But why is the Syntacticalist's position "unsatisfactory" if he has to adopt, say, an infinite sequence of metalanguages? Perhaps the answer is that the Syntacticalist's formal theories will be excessively cumbersome while the Non-Syntacticalist's will not.

We shall not pursue this point any further. Having noted our reservations, we shall agree with Montague that there is something unsatisfactory about the Syntacticalist's position if he has to adopt one of the suggested restrictions.

We should also note that Montague's claim that the Syntacticalist has certain difficulties which the Non-Syntacticalist does not have is far from obvious. As we pointed out in the previous section, the Non-Syntacticalist may have no difficulties within his formal theories but he may encounter difficulties, parallel to those of the Syntacticalist, in the metalanguage to which he is committed by his formal theories.

Let us turn to the evidence which Montague uses to support his Anti-Syntacticalist thesis. The evidence which he uses to support his thesis consists of an impor-

tant theorem which is proven in MONTAGUE (1963).

We shall introduce and explain some notation, which differs only slightly from Montague's, and then conclude this section by stating Montague's theorem.

Montague's theorem is about formal theories which contain theories of arithmetic as subtheories. We assume suitable Godel numberings to be imposed upon these theories. For any expression E of the theory $\underline{T}$, $\ulcorner nr(E) \urcorner$ will denote the numeral of $\underline{T}$ which, on the standard interpretation, denotes the Godel number of E. Where f is a number-theoretic function, $\ulcorner f(nr(E)) \urcorner$ will denote the numeral of $\underline{T}$ which denotes the value of f with the Godel number of E as argument.

"$\underline{Q}$" is a standard name for a theory of arithmetic also known as Robinson's arithmetic.[7] A characteristic feature of $\underline{Q}$ is that it is finitely axiomatizable, yet sufficient for the representation of all recursive functions. That $\underline{Q}$ is finitely axiomatizable will be seen to be essential to the proof of Montague's theorem since the conjunction of the arithmetic axioms must be able to appear as a sentence of the theory.

$\underline{Q}^{(\underline{Z})}$, referred to as "the relativization of $\underline{Q}$ with respect to $\underline{Z}$", is basically just $\underline{Q}$ with the predicate $\underline{Z}$

7. See MOSTOWSKI, ROBINSON and TARSKI (1953).

introduced into the appropriate places to be interpreted as "____ is a number". Thus the English "translation" of a typical sentence of $\underline{Q}$ might be

(4.2) For any x, there is a y such that y is the successor of x.

The relativization of this sentence in $\underline{Q}^{(\underline{Z})}$ would be rendered as

(4.3) For any x, if x is a number, then there is a y such that y is a number and y is the successor of x.

An advantage of $\underline{Q}^{(\underline{Z})}$ is that one may combine arithmetic with non-arithmetic axioms.

We now state the following result, which we refer to as Montague's Theorem. In the statement of the theorem "$\underline{N}$" may be any predicate, either a simple predicate letter or a complex matrix with one free variable.

(4.4) Montague's Theorem. If $\underline{T}$ is any theory such that

(1) $\underline{T}$ is an extension of $\underline{Q}$ (or of $\underline{Q}^{(\underline{Z})}$, for some one-place predicate $\underline{Z}$)

and there is a one-place predicate $\underline{N}$ of $\underline{T}$ such that for all sentences P, R of $\underline{T}$

(2) $\vdash_{\underline{T}} \underline{N}(nr(P)) \longrightarrow P$

(3) $\vdash_{\underline{T}} \underline{N}(nr(\underline{N}(nr(P)) \longrightarrow P))$

(4) $\vdash_{\underline{T}} \underline{N}(nr(P \longrightarrow R))$

$\longrightarrow (\underline{N}(nr(P)) \longrightarrow \underline{N}(nr(R)))$

(5) $\vdash_{\underline{T}} \underline{N}(nr(P))$, if P is a valid sentence of the first order functional calculus with identtity,

then $\underline{T}$ is inconsistent.[8]

8. MONTAGUE (1963), pp. 156-7.

/5/

A PROOF OF MONTAGUE'S THEOREM

In this section we shall present a proof for Montague's Theorem, (4.4). MONTAGUE (1963) gives an outline of the proof but it will be useful for us in what follows to have the ingredients of the proof made explicit.

An important ingredient in the Proof of Montague's Theorem is a principle to which Montague refers as "a principle of self reference".[1] It will be especially useful for us to have a proof of this principle which is stated, but not proven, in MONTAGUE (1963). The principle is embodied in the following result.

(5.1) If $\underline{T}$ is a theory in which all recursive functions are representable, F is a one-place predicate of $\underline{T}$, and f is a one-place recursive function; then there is a sentence S of $\underline{T}$ such that

$\vdash_{\underline{T}} S \leftrightarrow F(f(nr(S)))$.

Of a version of (5.1) in which "standard names" are used instead of Gödel numbers and in which f is the identity function, Kaplan and Montague say the following.

1. MONTAGUE (1963), p. 156.

> It has been shown by Godel that to provide for self-reference we need have at our disposal only the apparatus of elementary syntax. Then, whenever we are given a formula F whose sole free variable is 'x', we can find a sentence S which is provably equivalent to F($\overline{S}$), that is, the result of replacing in F the variable 'x' by the standard name of S. The sentence F($\overline{S}$) makes a certain assertion *about* the sentence S. Since S is provably equivalent to F($\overline{S}$), S makes the same assertion about S, and thus is self-referential. Besides this method and its variants, no other ways of treating self-referential sentences are known to us.[2]

To prove (5.1) we suppose that the antecedent conditions are fulfilled. Since the recursive function f is representable in T, there is a two-place predicate A(x,y) of T such that

(5.2) $\vdash_{\overline{T}}$ (y)(A(n,y) $\leftrightarrow$ y = k), if f(n) = k.

Let d be the following one-place recursive function. If m is the Godel number of a matrix of T whose only free variable is "x", then d(m) is the Godel number of the sentence of T which results from replacing all free occurrences of "x" in that matrix by the numeral of T which denotes m. If m is not the appropriate sort of Godel number, then d(m) = 0.

2. KAPLAN and MONTAGUE (1960), p. 82.

We let D(x,y) be a predicate of $\underline{T}$ which represents d, so that

(5.3) $\vdash_{\underline{T}}$ (y)(D(m,y) $\leftrightarrow$ y = n), if d(m) = n.

Consider the matrix

(5.4) (Ey)(Ez)(F(y) & A(z,y) & D(x,z)).

Suppose that this matrix has m as its Godel number and that the sentence

(5.5) (Ey)(Ez)(F(y) & A(z,y) & D(m,z))

has n as its Godel number, so that d(m) = n.

From (5.3) we infer

(5.6) $\vdash_{\underline{T}}$ (y)(D(m,y) $\leftrightarrow$ y = n).

Suppose that f(n) = k. Then from (5.2) we infer

(5.7) $\vdash_{\underline{T}}$ (y)(A(n,y) $\leftrightarrow$ y = k).

It is now sufficient to prove

(5.8) (Ey)(Ez)(F(y) & A(z,y) & D(m,z)) $\leftrightarrow$ F(k).

We will first give a proof of (5.8) from left to right. The proof will be a sequence of formulas each of which could be justified in the first order functional calculus with identity.

(5.9) 1. $(Ey)(Ez)(F(y) \& A(z,y) \& D(m,z))$

2. $F(a) \& A(b,a) \& D(m,b)$

3. $b = n$ from 2, (5.6)

4. $A(n,a)$ from 2, 3

5. $a = k$ from 4, (5.7)

6. $F(k)$ from 2, 5

The proof of (5.8) from left to right can be as follows.

(5.10) 1. $F(k)$

2. $A(n,k)$ from (5.7)

3. $D(m,n)$ from (5.6)

4. $F(k) \& A(n,k) \& D(m,n)$

5. $(Ey)(Ez)(F(y) \& A(z,y) \& D(m,z))$

To prove Montague's Theorem we assume that the antecedent conditions of (4.4) are fulfilled. Let A be the conjunction of the axioms of $\underline{Q}$, or of $\underline{Q}^{(\underline{Z})}$. A is sufficient for the derivation of an equivalence of the form of (5.2) for any recursive function. It is clear that there is a recursive function which maps the Godel number of a sentence of $\underline{T}$ to the Godel number of the conditional whose antecedent is A and whose consequent is the negation of the original sentence. By (5.1), there is a sentence S of $\underline{T}$ such that

(5.11) $\vdash_{\underline{T}} S \longleftrightarrow \underline{N}(nr(A \longrightarrow -S))$.

(5.11) is a number theoretic truth, which means that

(5.12) $A \longrightarrow (S \longleftrightarrow \underline{N}(nr(A \longrightarrow -S)))$

is a valid sentence of the first order functional calculus with identity. Since (5.12) truth-functionally implies

(5.13) $(\underline{N}(nr(A \longrightarrow -S)) \longrightarrow (A \longrightarrow -S)) \longrightarrow (A \longrightarrow -S)$

this latter sentence is also valid.

From (5.13), by condition (5) of (4.4), we may infer

(5.14) $\vdash_{\underline{T}} \underline{N}(nr((\underline{N}(nr(A \longrightarrow -S)) \longrightarrow (A \longrightarrow -S)) \longrightarrow (A \longrightarrow -S)))$.

By condition (3) of (4.4)

(5.15) $\vdash_{\underline{T}} \underline{N}(nr(\underline{N}(nr(A \longrightarrow -S)) \longrightarrow (A \longrightarrow -S)))$.

From (5.14), (5.15), and the relevant instance of condition (4) of (4.4) we may infer by truth-functional logic

(5.16) $\vdash_{\underline{T}} \underline{N}(nr(A \longrightarrow -S))$.

From (5.16) and the relevant instance of condition (2) of (4.4) we obtain

(5.17) $\vdash_{\underline{T}} A \longrightarrow -S$.

Since the axioms of $\underline{Q}$, or $\underline{Q}^{(\underline{Z})}$, are theorems of $\underline{T}$ we have $\vdash_{\underline{T}}$ A, and hence from (5.17)

(5.18) $\vdash_{\underline{T}}$ -S.

But from (5.11) and (5.16) by truth-functional logic

(5.19) $\vdash_{\underline{T}}$ S.

Thus $\underline{T}$ is inconsistent.

It should be noted that at no point in the preceding proof was it assumed that A, the conjunction of the axioms of Q, was itself necessary. All that was required to infer (5.18) from (5.17) was that A be true.

/6/

THE "PARITY OF NAMING" THESIS

In this section we shall begin to discuss the significance of Montague's Theorem, (4.4). Our interests here are twofold. First, we want to know what relevance this theorem has to our original question about the existence of a modal antinomy. Second, we want to examine Montague's claim that his theorem shows that the attempt to maintain a syntactical theory of modality at least as strong as Lewis' S1 is doomed to failure.

Our first interest can be served by pointing out the similarity between the antinomy of the liar and the customary proof of the theorem known as Tarski's Theorem. This theorem may be stated as follows.

(6.1) Tarski's Theorem. If $\underline{T}$ is any theory such that

(1) $\underline{T}$ is an extension of $\underline{Q}$ (or of $\underline{Q}^{(\underline{Z})}$ for some one-place predicate $\underline{Z}$)

and there is a one-place predicate $\underline{Tr}$ of $\underline{T}$, such that for any sentence P of $\underline{T}$

(2) $\vdash_{\underline{T}} \underline{Tr}(nr(P)) \leftrightarrow P$,

then $\underline{T}$ is inconsistent.

The customary proof of (6.1) involves the construc-

tion of a sentence which "says" that it, itself, is not true. That is, a sentence S such that

(6.2) $\vdash_{\underline{T}} S \leftrightarrow \underline{Tr}(nr(\neg S))$.

S is readily seen to be a formal representation of the sentence used in Section One to generate the antinomy of the liar.

By making the statement of (6.1) parallel to the statement of (4.4) we create a *prima facie* case for the supposition that a modal antinomy can be constructed by translating the proof of (4.4) into ordinary English in the same way that the antinomy of the liar can be constructed by translating the proof of (6.1) into ordinary English. This shows how Montague's Theorem is relevant to the question of the existence of a modal antinomy.

The question of translatability is a crucial aspect of both of the questions to be examined in this section. We have just seen that the possibility of making a translation of the proof of (4.4) into ordinary English must be justified in order to use (4.4) to argue for the existence of a modal antinomy. But also, in connection with our second question, the possibility of making a translation of the proof of (4.4) into other formal languages must be justified in order to use (4.4) to argue for Montague's Anti-Syntacticalist Thesis.

The proof of (4.4) makes it clear that any formal theory which meets the antecedent conditions of (4.4) cannot be used as a formalization of a syntactical theory of modality. This still leaves open the possibility that some other formal theory, one which did not meet the antecedent conditions of (4.4), might be a suitable formalization of a syntactical theory of modality.

In order to support his Anti-Syntacticalist Thesis it is incumbent upon Montague to show that there is no such alternative theory. To show that there is no such alternative theory it must be shown that any other formal theory is such that either a construction like that used in the proof of (4.4) can be carried out in that theory or else that the theory is in other respects inadequate. That is, it must be shown that the antecedent conditions of (4.4) are such that it would be reasonable to expect any proposed syntactical theory of modality to meet them.

When we look at the proof of (6.1) and its translation into ordinary English it soon becomes apparent that what the question of translatability amounts to is the question of translatability of alternative naming conventions. The self referential sentence (6.2) seems to differ from the self referential sentence used in Section One only in the method of naming sentences. In (6.2) a sentence is named by its Godel number. In Section One a

sentence was named by enclosing it in quotation marks. In the case of the antinomy of the liar we see that the antinomy can be constructed no matter which naming convention is adopted.

In Section Three, where we sketched out certain features of a language L suitable for expressing syntactical modal statements, we provided L with apparatus for naming sentences in a way analogous to the use of quotation marks in ordinary English. If Montague's Anti-Syntacticalist Thesis is to be plausible, he must argue that a theorem analogous to (4.4) can be proven for a formal theory which chooses to use such an alternative naming convention.

Montague does, in fact, so argue, as can be seen in the following quotations.

> ... we suppose syntax to be arithmetized and use as syntax languages those languages which contain the symbols of arithmetic ...[1]
>
> We can associate with each expression a term of Q which can be regarded as the standard name of that expression; to be specific we associate with the expression s the name "nr(s)" ...[2]
>
> This approach is by no means essential, and

1. MONTAGUE (1963), p. 158.
2. Ibid, p. 156. The notation for Godel numbering has been altered slightly.

> is adopted only to allow us to build on terminology and results already present in the literature. An equivalent and perhaps more natural approach would employ a syntax language ... which speaks directly about expressions.[3]

The gist of Montague's claim here is that Godel numbering is only one of several possible ways of naming expressions and that equivalent results can be obtained if alternative naming conventions are adopted. We shall refer to this claim as the Parity of Naming Thesis and the preceding statement of the claim as the vague form of the thesis.

Another explicit endorsement of the Parity of Naming Thesis can be found in KAPLAN and MONTAGUE (1960).

> It is ... desirable to introduce a system of names of expressions. Thus if $\underline{E}$ is any expression, $\overline{\underline{E}}$ is to be the <u>standard name</u> of $\underline{E}$, constructed according to one of several alternative conventions. We might, for instance, construe $\overline{\underline{E}}$ as the result of enclosing $\underline{E}$ in quotes. Within technical literature a more common practice is to identify $\overline{\underline{E}}$ with the numeral corresponding to the Godel-number of $\underline{E}$. As a third alternative, we could regard $\overline{\underline{E}}$ as the structural-descriptive name of $\underline{E}$ (within some well-determined metamathematical theory). A foundation for our later argu-

3. <u>Ibid</u>, p. 158n.

ments could be erected on the basis of any one of these conventions.[4]

We see that Montague is committed to some form of the Parity of Naming Thesis in order to support his Anti-Syntacticalist Thesis. We also see that some form of the thesis must be true if we are to construe the proof of Montague's Theorem as establishing the existence of a modal antinomy.

The Parity of Naming Thesis in its vague form is indeed far from clear. We shall conclude this section with a precise preliminary formulation of the thesis. As we shall see, there is good reason to think that Montague is not committed to the strong form of the thesis which we shall state in this section. It will be useful, however, to have this preliminary formulation of the thesis for our later discussion.

Our aim is to state the thesis in such a way that it will have some implications for a theory in L, the language whose general features were described in Section Three. Consider then the following claim.

(6.3) The Parity of Naming Thesis (strong form). If T is a theory in L and there is a one-place predicate N of T, such that for all sentences P, R

4. KAPLAN and MONTAGUE (1960), p. 80.

of $\underline{T}$

(1) $\vdash_{\underline{T}}$ $\underline{N}(\text{'P'}) \longrightarrow P$

(2) $\vdash_{\underline{T}}$ $\underline{N}(\text{'}\underline{N}(\text{'P'}) \longrightarrow P\text{'})$

(3) $\vdash_{\underline{T}}$ $\underline{N}(\text{'P} \longrightarrow R\text{'}) \longrightarrow (\underline{N}(\text{'P'}) \longrightarrow \underline{N}(\text{'R'}))$

(4) $\vdash_{\underline{T}}$ $\underline{N}(\text{'P'})$, if P is a valid sentence of the first order functional calculus with identity,

then $\underline{T}$ is inconsistent.

Strictly speaking the strong form of the Parity of Naming Thesis should be the claim that (4.4) entails (6.3). But since (4.4) is true this entailment claim is equivalent to the simple claim that (6.3) is true.

/7/

THE LANGUAGE L AND THE THEORY SM

In this section we shall give an exact characterization of the language L whose general features were discussed in Section Three. We shall then construct a formal theory in L which may serve as a formalization of a syntactical theory of modality. This theory will be called SM.

(7.1) The language L contains the following symbols:[1]

(1) "v", "-", "'", "E", "(", and ")",

(2) an infinite number of individual constants; including "a", "b", "c", etc.,

(3) an infinite number of individual variables; including "x", "y", "z", etc.,

(4) for each n greater than zero, an infinite number of n-place predicate symbols; "N" and "=" being among the one and two-place predicate symbols respectively,

(5) for each n greater than zero, an infinite number of n-place function symbols; "neg" and "disj" being among the one and two-place function symbols respectively.

1. We let $\ulcorner P \longrightarrow Q \urcorner$ abbreviate $\ulcorner -P \text{ v } Q \urcorner$, and $\ulcorner (v)P \urcorner$ abbreviate $\ulcorner -(Ev)-P \urcorner$. Obvious conventions for omitting parentheses and introducing commas are adopted.

We cannot specify the terms of L independently of the formulas of L since the single quotation marks are, in effect, a function for forming terms from sentences. What are terms of L will depend on what are sentences of L, but the sentences of L cannot be specified without mentioning the terms of L. Thus (7.2) and (7.3) must be taken as interdependent.

(7.2) The formulas of L include all and only those expressions so specified by the following four clauses:

(1) $\ulcorner G(t_1 \ldots t_n) \urcorner$ is a formula if G is an n-place predicate symbol and each of $t_1, .., t_n$ are terms,

(2) $\ulcorner -P \urcorner$ is a formula if P is a formula,

(3) $\ulcorner P \text{ v } Q \urcorner$ is a formula if P and Q are formulas,

(4) $\ulcorner (Ev)P \urcorner$ is a formula if P is a formula and v is a variable.

Formulas of the form $\ulcorner =(t_1 t_2) \urcorner$ will be written $\ulcorner t_1 = t_2 \urcorner$.

An occurrence of a variable v in a formula P is <u>free</u> just in case that occurrence of v is not in any part of P of the form $\ulcorner (Ev)Q \urcorner$ where Q is a formula. An occurrence of a variable which is not a free occurrence is a <u>bound</u> occurrence. The <u>sentences</u> of L are just those formulas of L which have no free occurrences of variables.

(7.3) The terms of $\underline{L}$ include all and only those expressions so specified by the following three clauses:

(1) all individual constants and all variables are terms,

(2) $\ulcorner f(t_1 \ldots t_n) \urcorner$ is a term if f is an n-place function symbol and each of $t_1, .., t_n$ are terms,

(3) $\ulcorner \text{'P'} \urcorner$ is a term if P is a sentence.

It should be noted that (7.2) and (7.3) allow expressions of the form $\ulcorner \text{'P'} \urcorner$ to be terms only when P is a sentence.

We have already specified which occurrences of variables are bound occurrences. As we indicated in Section Three we intend to expand the bound-free distinction to cover occurrences of terms other than variables. For our purposes it is convenient to expand the distinction so that it applies to occurrences of symbols which may not even be terms.

We will say that an occurrence of a symbol, or of a sequence of symbols, s in a formula P is bound if that occurrence of s either (1) contains an occurrence of a variable such that that occurrence of the variable is bound in P, or (2) s occurs in a part of P of the form $\ulcorner \text{'Q'} \urcorner$ where Q is a formula and s is not identical with $\ulcorner \text{'Q'} \urcorner$.

We will consider a theory in L to be determined by (1) the specification of a set of individual constants and a set of predicate and function symbols, which set is to include "=", (2) the specification of a set of axioms, and (3) the specification of rules of inference.

It is convenient to have a set of axiom schemata and rules of inference which shall be common to every theory in L. Instances of these common axiom schemata and rules of inference will be known as a theory's logical axioms and logical inferences respectively. The remaining axioms and rules of inference for a theory will be known as that theory's non-logical axioms and non-logical inferences respectively.

The purpose of the shared logical axioms and logical inferences is to guarantee (1) that every logically valid formula of the theory is provable in the theory, and (2) that every formula of the theory which is a logical consequence of non-logical axioms is provable in the theory.

There are many sets of logical axioms and rules of inference available in the literature, any of which would be suitable for our purposes. We choose the following.

(7.4) The logical axioms of a theory in L are the instances of the following schemata; where P, Q, and R are formulas of the theory; v is a variable; t and u are terms of the theory; P(t//u)

results from P by replacing <u>all</u> free occurrences of t in P by u if none of the new occurrences of u will be bound, otherwise P(t//u) is P; and P(t/u) results from P by replacing <u>at most one</u> free occurrence of t in P by u, with the proviso that the new occurrence of u is not bound:

(1) $P \longrightarrow (Q \longrightarrow P)$

(2) $(P \longrightarrow (Q \longrightarrow R)) \longrightarrow ((P \longrightarrow Q) \longrightarrow (P \longrightarrow R))$

(3) $(-P \longrightarrow -Q) \longrightarrow ((-P \longrightarrow Q) \longrightarrow P)$

(4) $(v)P \longrightarrow P(v//t)$

(5) $(v)(P \longrightarrow Q) \longrightarrow (P \longrightarrow (v)Q)$, provided that v does not occur free in P,

(6) $t = t$

(7) $t = u \longrightarrow (P \longrightarrow P(t/u))$.

(7.5) The logical inferences of a theory in <u>L</u> are to be the following, where P and Q are formulas of the theory and v is a variable:

(1) from P and $\ulcorner P \longrightarrow Q \urcorner$ to infer Q

(2) from P to infer $\ulcorner (v)P \urcorner$.

The theroy <u>SM</u> contains the one-place predicate symbol "N", the one-place function symbol "neg", and the two-place function symbol "disj". <u>SM</u> also contains an infinite number of individual constants, including "a", "b", "c", etc.

(7.6) The non-logical axioms of SM are the instances of the following schemata, where P and Q are sentences:

(1) $N('P') \longrightarrow P$

(2) $N('P \longrightarrow Q') \longrightarrow (N('P') \longrightarrow N('Q'))$

(3) $-N('P') \longrightarrow N('-N('P')')$

(4) $neg('P') = '-P'$

(5) $disj('P','Q') = 'P \vee Q'$

(6) $'P' \neq 'Q'$, if P and Q are different sentences.

The purpose of the function symbols "neg" and "disj", which are sufficient to represent all truth functions, is to allow us to represent modal predicates other than "____ is necessary" in SM. Without these function sym- we could represent a predicate such as "____ is possible" only in the cases where the argument expression was a term of the form $\ulcorner'Q'\urcorner$. That is, we could regard $\ulcorner P('Q')\urcorner$ as an abbreviation of $\ulcorner -N('-Q')\urcorner$.

In the general case, where t might be a term which is not of the form $\ulcorner'Q'\urcorner$, we need the function symbol "neg" so that we might regard $\ulcorner P(t)\urcorner$ as an abbreviation of $\ulcorner -N(neg(t))\urcorner$. In virtue of axiom (4) of (7.6) we see that the latter abbreviation yields the former as a consequence when t is of the form $\ulcorner'Q'\urcorner$.

(7.7) The only non-logical inference in SM is the following, where P is a sentence:

(1) from P to infer $\ulcorner N(\text{'P'})\urcorner$.

As a syntactical theory of modality SM contains as theorems the syntactical analogues of all theorems of Lewis' system S5. Axioms (1) through (3) of (7.6), together with the non-logical inference (7.7), are just the syntactical analogues of the common axiomatization of S5.

We chose SM as a formalization of a syntactical theory of modality, even though it is a stronger theory than many Syntacticalists would care to accept, because our goal is to prove the consistency of SM. The consistency of all weaker theories will follow from the consistency of SM.

/8/

A SPECIFICATION OF A METHOD FOR PROVING THE CONSISTENCY OF SM

In this section we shall describe the method which we intend to use to prove the consistency of SM. In the next section we shall carry out the proof. Before describing the method to be used, we discuss certain problems raised by various attempts to construct a consistency proof for SM. This discussion, which is basically a listing of tempting, but ultimately fruitless, methods for proving the consistency of SM, may be viewed as a way of sharing some of the contents of our waste basket.

There are three obvious suggestions on how to prove the consistency of SM. The first of these is to find a mapping from the formulas of SM into the formulas of a theory T whose consistency no one doubts. It would have to be shown that this mapping maps all theorems of SM into theorems of T and maps at least one formula of SM into a non-theorem of T.

The second suggestion is to construct a set of many-valued matrices, setting certain values aside as designated values. Then it would have to be shown that each axiom of SM had one of the designated values and that the rules of inference yielded formulas with a designated

value when applied to formulas with a designated value. The proof would be completed by finding a formula of SM which did not have a designated value.

The third suggestion is to find a model such that the axioms of SM are all true in the model and the rules of inference yield formulas true in the model when applied to formulas true in the model.

The difficulty with the first suggestion is that there is no obvious theory T which has a feature analogous to SM's use of single quotation marks to form terms from sentences.

It might be thought that the formulas of SM could be mapped into the formulas of a non-syntactical theory of modality in some natural manner. Identity statements of SM would presumably map into equivalence statements of some sort. But these equivalence statements could not be material equivalences, since the natural analogue of

(8.1) $a = b \longrightarrow (N(a) \longrightarrow N(b))$

would be

(8.2) $(A \longleftrightarrow B) \longrightarrow (\Box A \longrightarrow \Box B)$.

(8.1) is a theorem of SM but (8.2) is not a theorem of any non-syntactical theory of modality. Perhaps this serves to show, not that there is no mapping of the de-

sired sort, but only that identity statements should not be mapped into material equivalences.

If identity statements were mapped into logical equivalences (8.1) would probably be mapped into

(8.3) $\Box (A \leftrightarrow B) \rightarrow (\Box A \rightarrow \Box B)$

which is a recognized non-syntactical theorem. But certain difficulties are also raised by the suggestion that identity statements be mapped into logical equivalences. Consider

(8.4) 'N(a)' $\neq$ '- -N(a)'.

Its non-syntactical analogue cannot be

(8.5) $- \Box (\Box A \rightarrow - - \Box A)$

since the latter is the negation of a non-syntactical modal theorem.

Consider the second suggestion. Suppose that we have 1,..,n as values in some many-valued system. We can probably always construct sentences $A_1,..,A_n$ such that A_i has the value i, $1 \leq i \leq n$. A natural way to interpret $\ulcorner$'x = 'P'$\urcorner$ is as having a designated value just in case "x" and $\ulcorner$'P'$\urcorner$ have the same value, while $\ulcorner$'P' = 'Q'$\urcorner$ is naturally interpreted as having a designated value just in case P and Q are the same sentence.

Consider the following instance of axiom (4) of (7.6).

$$(8.6)\quad (x)(\, x = \text{'}A_1\text{'} \vee \ldots \vee x = \text{'}A_n\text{'}\,) \longrightarrow \text{'}B\text{'} = \text{'}A_1\text{'} \vee \ldots \vee \text{'}B\text{'} = \text{'}A_n\text{'}$$

where B is different from each A_i.

The antecedent of (8.6) will probably have a designated value and its consequent an undesignated value. A conditional with antecedent of designated value and consequent of undesignated value must be assigned an undesignated value. Otherwise we could not show that *modus ponens* yields only formulas with a designated value when applied to formulas with a designated value. Hence (8.6) is an example of a theorem of SM which will probably have an undesignated value.

Perhaps (8.6) only shows that sentences of the form $\ulcorner \text{'}P\text{'} = \text{'}Q\text{'} \urcorner$ should not be assigned a designated value in just those cases where P and Q are the same sentence. Perhaps $\ulcorner \text{'}P\text{'} = \text{'}Q\text{'} \urcorner$ should be assigned a designated value whenever P and Q are assigned the same value.

The following difficulty is raised by this suggestion. We can easily prove any instance of

$$(8.7)\quad \text{'}P\text{'} = \text{'}Q\text{'} \longrightarrow (N(\text{'}P \longrightarrow P\text{'}) \longrightarrow N(\text{'}P \longrightarrow Q\text{'}))$$

in SM. Presumably $\ulcorner N(\text{'}P \longrightarrow P\text{'}) \urcorner$ will always have a de-

signated value while $\ulcorner N(\text{'}P \longrightarrow Q\text{'})\urcorner$ need not.

Our third suggestion was to look for a model for the theory SM. A natural choice for a model is the set of sentences of SM itself. P could always be assigned to the term $\ulcorner\text{'}P\text{'}\urcorner$ while different interpretations could assign different sentences to the individual symbols of SM. Here we could declare that $\ulcorner N(\text{'}a = \text{'}P\text{''})\urcorner$ is false since there is only one interpretation, the one in which P is assigned to "a", which would make $\ulcorner a = \text{'}P\text{'}\urcorner$ true.

Difficulties appear when we consider expressions such as

(8.8) N(disj(a,'P')).

The truth of this expression will presumably depend on whether a true sentence is assigned to $\ulcorner disj(a,\text{'}P\text{'})\urcorner$ by all interpretations. But under one interpretation (8.8) itself is assigned to "a", which means that we have to determine the truth value of

(8.9) N(disj(a,'P')) v P.

It is not unreasonable to suppose that we might have to know the truth value of (8.8) in order to determine the truth value of (8.9). We are thus left in the embarassing position of having to already know the truth value of (8.8) in order to determine the truth value of (8.8).

A similar problem arises with universally quantified statements since one of the instances of $\ulcorner(v)P\urcorner$ is $P(v//'(v)P')$.

The method which we finally hit upon to prove the consistency of SM is a syntactical reduction procedure. This procedure is an adaptation of the method of proof used in "tree" systems. These systems are quite common in the literature but have, unfortunately, never gained much popularity.[1]

In these systems trees are constructed for formulas. Trees which satisfy certain requirements are said to be closed. We will show that every axiom of SM has a closed tree, that the rules of inference of SM yield formulas with closed trees when applied to formulas with closed trees, and, finally, that there is at least one formula of SM which does not have a closed tree. This will be sufficient to show the consistency of SM.

A tree may be deined abstractly as a set of points which obey the following conditions.

1. The basic principles of the tree systems can be traced back to Herbrand, Gentzen, and Beth. The closest relative of the system that we shall use is that of ANDERSON and BELKNAP (1959). Further references can be found in SMULLYAN (1968). All of these systems are designed for the first order functional calculus (usually not with identity). The adaptation to a syntactical theory of modality, though fairly obvious, is ours.

There is a function f from the set of points into the positive integers, f(p) being the level of the point p. There is to be exactly one point of level 1, this point being the base of the tree. There is a "successor" relation S(p,q) defined on the points. S(p,q) is to hold just in case $f(q) = f(p) + 1$. A further condition is that for every point q such that $f(q) \neq 1$, there is exactly one point p such that S(p,q) holds.

A tree for a formula P is a tree which has a formula at every point of the tree, P being at the base of the tree. A formula which is placed at a point which has no successor is a terminal formula of the tree.

Certain formulas are designated as tree axioms.

(8.10) A formula P is a tree axiom if and only if it meets one of the following three conditions, where Q is a formula, R and S are any two different sentences, and t is a term:

(1) both Q and $\ulcorner\neg Q\urcorner$ are disjuncts of P,

(2) $\ulcorner t = t\urcorner$ is a disjunct of P,

(3) $\ulcorner \text{'R'} \neq \text{'S'}\urcorner$ is a disjunct of P.

We say that a tree for a formula is a closed tree just in case it meets the following three conditions. (1) The tree has only a finite number of points, (2) every terminal formula is a tree axiom, and (3) every non-terminal formula and its successor(s) stand in the

relation of conclusion to premise(s) of one of the rules R1 to R16 given below in (8.11).

The rules R1 to R16 are presented as schemata in which Q may be void. The upper lines are the premises of the rules and the lower lines the conclusions.

The terminology of our system is a bit odd in that in working upwards we work "backwards" from conclusion to premises. However, our terminology is not quite as odd as that of many tree systems, for example, those where trees "grow" downwards.

It is convenient to adopt two further conventions. First we stipulate that the rules may be applied only to formulas which are not tree axioms. Under this stipulation only a terminal formula may be a tree axiom.

Second, if X* and Y* result from X and Y respectively by reordering and/or regrouping disjuncts, and if X and Y are related as premise and conclusion of one of our rules; then X* and Y* will also be counted as an instance of a premise and conclusion of that rule.

(8.11) R1:

$$\frac{P \vee Q}{\neg\neg P \vee Q}$$

R2:

$$\frac{\neg P \vee Q \qquad \neg R \vee Q}{\neg(P \vee R) \vee Q}$$

R3: $-P(v//v^*) \vee Q$ / $-(Ev)P \vee Q$ — where v^* is a variable which has no free occurrences in the conclusion

R4: $(Ev)P \vee P(v//t) \vee Q$ / $(Ev)P \vee Q$

R5: $t \neq u \vee P(t/u)$ / $t \neq u \vee P$

R6: $u \neq t \vee Q$ / $t \neq u \vee Q$

R7: $a = a \vee Q$ / $N('P') \vee Q$ — if P has a closed tree

$a \neq a \vee Q$ / $N('P') \vee Q$ — if P has a closed tree

R8: $a \neq a \vee Q$ / $-N('P') \vee Q$ — if P has a closed tree

$a = a \vee Q$ / $-N('P') \vee Q$ — if P has no closed tree

R9: $P(neg('R')/'-R')$ / P

R10: $P('-R'/neg('R'))$ / P

R11: $P(disj('R','S')/'R \vee S')$ / P

R12: P('R v S'/disj('R','S'))

P

R13: $t \neq u \vee t \neq u \vee Q$

$t \neq u \vee Q$

R14: $-N(t) \vee -N(t) \vee Q$

$-N(t) \vee Q$

R15: $-(Ev)P \vee -(Ev)P \vee Q$

$-(Ev)P \vee Q$

R16: $P \vee t \neq t$

P

We refer to R5, R9, R10, R11, and R12 as substitution rules. In the case of the substitution rules we do not say that any disjunct of the premise has been "operated upon". In the case of a non-substitution rule, other than R16, we say that the disjunct of the premise other than Q is the disjunct that has been operated upon. In the case of R16 we say that the rightmost disjunct of the premise has been operated upon.

Tree systems such as that of ANDERSON and BELKNAP (1959) come equipped with a systematic technique for constructing trees. Indeed one of the principal benefits of using their system is their deterministic search procedure for closed trees. That is, if a formula has a closed tree they guarantee that one will be able to con-

struct it. In their system we may speak of the tree for a formula since each formula has a unique tree.

We have not given a systematic technique for the application of our rules. The rules may be applied in any arbitrary order. In our system a formula with a closed tree will, in general, also have many trees that are not closed.

R7 and R8 are not constructive rules. It might be felt that a systematic technique for tree construction would be an advantage in applying these rules. While it is true that we could make slight alterations in our rules to formulate a technique that would enable us to determine in a finite number of steps whether or not a formula of SM had a closed tree, we do not bother to do so because our main interest is in proving the consistency of SM and not in constructing trees.

It should be kept in mind that such a systematic technique would not provide a decision procedure for SM since there are non-theorems which have closed trees. $\ulcorner \sim N(\text{'a = b'}) \urcorner$, for example, has a closed tree but is probably not a theorem of SM. The use of trees is a convenient device for proving the consistency of SM. One has not discovered anything of the slightest philosophical interest about a formula of SM when he discovers that it has a closed tree.

/9/

A PROOF THAT THE CUT-ELIMINATION THEOREM FOR FORMULAS OF SM IMPLIES THE CONSISTENCY OF SM

In this section we shall prove, subject to a certain hypothesis that every axiom of SM has a closed tree, that the rules of inference yield formulas with closed trees when applied to formulas with closed trees, and that there is at least one formula of SM which does not have a closed tree. Granted the hypothesis, this will be sufficient to prove the consistency of SM.

As indicated by the title of this section, the hypothesis in question is the Cut-Elimination Theorem for formulas of SM. For organizational convenience we defer the proof of this theorem to the following section.

We first show that each logical axiom of SM has a closed tree.

When axiom (1) of (7.4) is written in primitive notation and associative parentheses are ignored, the result is

(9.1) $-P \vee -Q \vee P$

which is a tree axiom.

Axiom (2) of (7.4) is

(9.2) -(-P v -Q v R) v -(-P v Q) v -P v R.

An application of R2 to (9.2), operating upon the second disjunct, yields formulas which can be written as

(9.3) -(-P v -Q v R) v - -P v -P v R

and

(9.4) -(-P v -Q v R) v -P v -Q v R

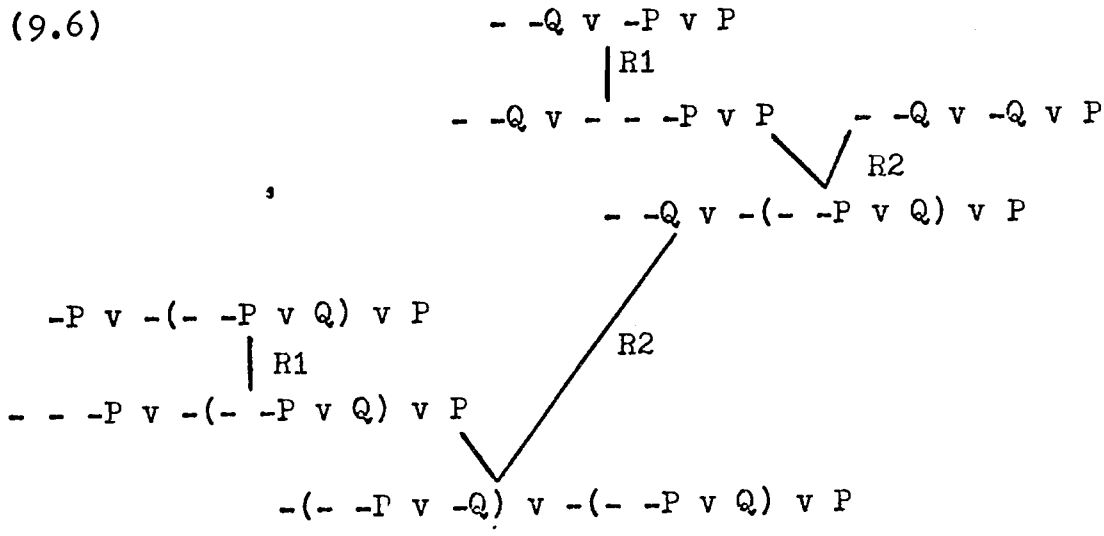
both of which are tree axioms.

Axiom (3) of (7.4) is

(9.5) -(- -P v -Q) v -(- -P v Q) v P.

The following is a closed tree for (9.5).

(9.6)



Axiom (4) of (7.4) is

(9.7) - -(Ev)-P v P(v//t).

An application of R1 to (9.7) yields

(9.8) (Ev)-P v P(v//t)

and an application of R4 to (9.8) yields

(9.9) (Ev)-P v -P(v//t) v P(v//t)

which is a tree axiom.

Axiom (5) of (7.4) is

(9.10) - -(Ev)-(-P v Q) v -P v -(Ev)-Q

where v does not occur free in Q.

The following is a closed tree for (9.10).

(9.11) (Ev)-(-P v Q) v - -P v -P v Q(v//v*)

R2 (Ev)-(-Pv Q) v -Q(v//v*) v -P v - -Q(v//v*)

(Ev)-(-P v Q) v -(-P v Q(v//v*)) v -P v - -Q(v//v*)

|R4

(Ev)-(-P v Q) v -P v - -Q(v//v*)

|R3

(Ev)-(-P v Q) v -P v -(Ev)-Q

|R1

- -(Ev)-(-P v Q) v -P v -(Ev)-Q

Axiom (6) of (7.4) is

(9.12) t = t

which is a tree axiom.

Axiom (7) of (7.4) is

(9.13) $t \neq u \vee -P \vee P(t/u)$.

An application of R5 to (9.13) yields

(9.14) $t \neq u \vee -P(t/u) \vee P(t/u)$

which is a tree axiom. This completes our proof that every logical axiom of SM has a closed tree.

In the case of the non-logical axioms of SM some argument is required. Axiom (1) of (7.6) is

(9.15) $-N('P') \vee P$.

If P does not have a closed tree then, by R8, the successor of (9.15) may be

(9.16) $a = a \vee P$

which is a tree axiom.

If P does have a closed tree, then a closed tree for (9.15) may be constructed by simply adding $\ulcorner -N('P') \urcorner$ as a disjunct to each formula in the closed tree for P.

In either case, then, there is a closed tree for (9.15).

Axiom (2) of (7.6) is

(9.16) $-N('P \longrightarrow Q') \vee -N('P') \vee N('Q')$.

We observe that (9.16) has a closed tree if either $\ulcorner P \longrightarrow Q \urcorner$ or P fails to have a closed tree. The only possibly difficult case, then, is when $\ulcorner P \longrightarrow Q \urcorner$ and P both have closed trees. If Q has a closed tree in this case then (9.16) will have a closed tree. Thus, the problem of showing that (9.16) has a closed tree reduces to the problem of showing that our logical inference modus ponens preserves tree closure.

Axiom (3) of (7.6) is

(9.17) - -N('P') v N('-N('P')')

which has a closed tree if

(9.18) N('P') v N('-N('P')')

has a closed tree.

Suppose that P has a closed tree. Then, by R7, the successor of (9.18) may be

(9.20) a = a v N('-N('P')')

which is a tree axiom.

Suppose that P does not have a closed tree. Then there is a closed tree for $\ulcorner$-N('P')$\urcorner$. By R7 the successor of (9.18) may be

(9.21) N('P') v a = a

which is a tree axiom. In either case there is a closed tree for (9.18).

Axioms (4) and (5) of (7.6) are

(9.22) neg('P') = '-P'

and

(9.23) disj('P','Q') = 'P v Q'

respectively. Applying R9 and R11 gives us

(9.24) '-P' = '-P'

and

(9.25) 'P v Q' = 'P v Q'

as the respective successors of these two axioms.

Instances of axiom (6) of (7.6) are of the form

(9.26) 'P' ≠ 'Q'

where P and Q are different sentences. Every such instance is a tree axiom.

This completes our proof that all of the axioms of SM have closed trees. We turn now to the rules of inference for SM.

The non-logical inference (7.7) quite obviously preserves tree closure. If P has a closed tree, so does

$\ulcorner N(\text{'}P\text{'})\urcorner$.

Only the logical inferences remain. It is quite easy to show that universal generalization, rule (2) of (7.5), preserves tree closure. Suppose that P has a closed tree. P may be represented as P(v//v*) where v* is just v. Since v has no free occurrences in $\ulcorner(v)P\urcorner$, that is, in $\ulcorner\text{-}(Ev)\text{-}P\urcorner$, we may first use R1 to subjoin $\ulcorner\text{-}\ \text{-}P(v//v^*)\urcorner$ to the base of the closed tree for P and then use R3 to subjoin $\ulcorner\text{-}(Ev)\text{-}P\urcorner$.

What is most difficult has been left for last. We must show that modus ponens preserves tree closure.

Instead of dealing with modus ponens directly, it is more convenient to imagine that SM contains the following two rules of inference in place of modus ponens.

(9.27) Addition Rule. From P to infer $\ulcorner P \vee Q\urcorner$.

(9.28) Cut Rule. From $\ulcorner\text{-}P \vee Q\urcorner$ and $\ulcorner P \vee Q\urcorner$ to infer Q.

It is obvious that a system which has the Cut Rule and the Addition Rule will have modus ponens as a derived rule. It is also obvious that the Addition Rule preserves tree closure.[1]

We will have established that the rules of inference

1. This claim will be elaborated upon in footnote 2 of Section Ten.

for SM preserve tree closure once we prove the following.

(9.29) Cut-Elimination Theorem. If there are closed trees for both ⌜-P v Q⌝ and ⌜P v Q⌝, then there is a closed tree for Q.

(9.29) is an extended version of Gentzen's famous Hauptsatz or Cut-Elimination Theorem. We will prove (9.29) in the next section.

To complete our proof of the consistency of SM we must find a formula of SM which does not have a closed tree. We claim the following.

(9.30) There is no closed tree for "a ≠ a".

A rigorous proof of (9.30) would proceed as follows. For any term t, we let t* represent a term that could be "derived" from t by some finite number (possibly zero) of substitutions made according to R9, R10, R11, or R12. We then say that a formula is recalcitrant if each of its disjuncts is of the form ⌜t ≠ t*⌝ for some term t.

We prove by induction that no recalcitrant formulas have closed trees. We first show that no recalcitrant formulas are tree axioms. Then we let m be the least number of formulas required to construct a closed tree for a recalcitrant formula. Suppose that P is a recalcitrant formula which has a closed tree with exactly m formulas in it. Inspection of the forms of our rules can

show that the successor of P is also recalcitrant. Removing P from the base of its closed tree leaves us with a closed tree for the successor of P, a recalcitrant formula, which contains only m-1 formulas. This contradicts our induction hypothesis.

(9.30) is established by pointing out that "a ≠ a" is recalcitrant.

/10/

A PROOF OF THE CUT-ELIMINATION THEOREM FOR FORMULAS OF SM

Our purpose in this section is to prove the following result.[1]

(10.1) Cut-Elimination Theorem. If there are closed trees for both $\ulcorner P \vee Q \urcorner$ and $\ulcorner -P \vee Q \urcorner$, then there is a closed tree for Q.

To prove (10.1) we need some additional terminology. By the weight of a tree we shall mean the number of points in the tree. By the degree of a formula we shall mean the number of free occurrences of quantifiers and sentential connectives in that formula.

We shall say that a formula P is n-eliminable if, for every formula Q, there is a closed tree for Q whenever there is a closed tree for $\ulcorner P \vee Q \urcorner$ of weight n or less and a closed tree for $\ulcorner -P \vee Q \urcorner$. We shall say that P is eliminable if P is n-eliminable for every n.

There are some preliminary results which must be established before (10.1) can be proven. The first of these is a principle which is too obvious to require proof, but

1. Although our proof is so tedious that no one would wish to claim credit for it, certain features of it were nonetheless suggested by MENDELSON (1964), appendix, and SMULLYAN (1968), chapter 12.

for which it is convenient to have a name.

(10.2) The Extra Disjunct Principle. If there is a closed tree for P of weight n, then there is a closed tree for ⌜P v Q⌝ whose weight is not greater than n.[2]

The next result which we need is the following principle.

(10.3) The Invertibility Principle. If Q* and Q are related as premise and conclusion of one of our rules and there is a closed tree for Q, then there is a closed tree for Q*.

(10.3) requires us to prove a certain result about each of our 16 rules. The method of proof will not be the same for each rule.

R4, R13, R14, R15, and R16 are immediately seen to be invertible by the Extra Disjunct Principle.

For the case of R5, R6, R9, R10, R11, and R12 there are obvious ways to subjoin formulas to the base of the

2. The proof of (10.2) is not quite as trivial as we have suggested. It is not sufficient just to add Q as a disjunct to each formula in the closed tree for P, since Q may contain free occurrences of a variable introduced into the closed tree for P by R3. We must show that we can always replace such variables by other new variables, none of which conflict with variables occurring free in Q.

closed tree for the conclusion of the rule in order to obtain a closed tree for the premise of the rule.

Given a closed tree for $\ulcorner t \neq u \vee P \urcorner$, for example, we may first, by R6, subjoin an instance of $\ulcorner u \neq t \vee P \urcorner$, then, by R5, we may subjoin an instance of $\ulcorner u \neq t \vee P(t/u) \urcorner$, and finally, by R6 again, we may subjoin an instance of $\ulcorner t \neq u \vee P(t/u) \urcorner$. This shows the invertibility of R5. The argument is similar for the five other rules mentioned.

For the case of R1 and R2 we can prove (10.3) by a simple induction. The proofs are similar so we consider only R1.

Suppose that there is a closed tree for $\ulcorner - -P \vee Q \urcorner$ of weight one. Then, either Q contains disjunct(s) which make Q a tree axiom, in which case $\ulcorner P \vee Q \urcorner$ is also a tree axiom, or else Q contains either $\ulcorner -P \urcorner$ or $\ulcorner - - -P \urcorner$ as a disjunct. If Q contains $\ulcorner -P \urcorner$ as a disjunct, then $\ulcorner P \vee Q \urcorner$ is a tree axiom. If Q contains $\ulcorner - - -P \urcorner$ as a disjunct, then R1 allows us to write as the successor of $\ulcorner P \vee Q \urcorner$ a formula which contains both P and $\ulcorner -P \urcorner$ as disjuncts.

Suppose that there is a closed tree for the formula of the form $\ulcorner P \vee Q \urcorner$ whenever there is a closed tree for the formula of the form $\ulcorner - -P \vee Q \urcorner$ of weight less than n.

Let T be a closed tree for $\ulcorner - -P \vee Q \urcorner$ of weight n.

Inspection of the forms of the rules shows that if the first rule used in T is any rule other than R1 where $\ulcorner - -P \urcorner$ is operated upon, R2, or a substitution rule where a substitution is made in $\ulcorner - -P \urcorner$; then the successor of $\ulcorner - -P \vee Q \urcorner$ in T may be written as $\ulcorner - -P \vee Q^* \urcorner$ where $\ulcorner P \vee Q^* \urcorner$ is also related to $\ulcorner P \vee Q \urcorner$ as premise to conclusion of one of our rules.

Removing $\ulcorner - -P \vee Q \urcorner$ from the base of T leaves us with a closed tree for $\ulcorner - -P \vee Q^* \urcorner$ of weight n-1. By the induction hypothesis there is a closed tree for $\ulcorner P \vee Q^* \urcorner$. To the base of this latter tree we may subjoin an instance of $\ulcorner P \vee Q \urcorner$.

If the first rule used in T is R1 where $\ulcorner - -P \urcorner$ is operated upon, then there is obviously a closed tree for $\ulcorner P \vee Q \urcorner$.

If the first rule used in T is R2, then the successors of $\ulcorner - -P \vee Q \urcorner$ in T may be written as $\ulcorner - -P \vee Q^* \urcorner$ and $\ulcorner - -P \vee Q^{**} \urcorner$ where $\ulcorner P \vee Q^* \urcorner$ and $\ulcorner P \vee Q^{**} \urcorner$ are also related to $\ulcorner P \vee Q \urcorner$ as premises to conclusion of R2. Removing $\ulcorner - -P \vee Q \urcorner$ from the base of T leaves us with closed trees for $\ulcorner - -P \vee Q^* \urcorner$ and $\ulcorner - -P \vee Q^{**} \urcorner$ each of weight less than n, By the induction hypothesis there are closed trees for $\ulcorner P \vee Q^* \urcorner$ and $\ulcorner P \vee Q^{**} \urcorner$. To the bases of these closed trees we may subjoin a common instance of $\ulcorner P \vee Q \urcorner$.

If the first rule used in T is a substitution rule where a substitution is made in $\ulcorner - -P\urcorner$, then the successor of $\ulcorner - -P \vee Q\urcorner$ in T may be written as $\ulcorner - -P^* \vee Q\urcorner$ where $\ulcorner P^* \vee Q\urcorner$ is also related to $\ulcorner P \vee Q\urcorner$ as premise to conclusion of one of our rules. Removing $\ulcorner - -P \vee Q\urcorner$ from the base of T leaves us with a closed tree for $\ulcorner - -P^* \vee Q\urcorner$ of weight n-1. By the induction hypothesis there is a closed tree for $\ulcorner P^* \vee Q\urcorner$. To the base of this latter tree we may subjoin an instance of $\ulcorner P \vee Q\urcorner$.

The invertibility of R3 follows from a more general result which, since we will need it later in our proof, we prove at this time. The following is the more general result.

(10.4) If there is a closed tree for $\ulcorner -(Ev)P \vee Q\urcorner$, then there is a closed tree for $\ulcorner -P(v//t) \vee Q\urcorner$ where t may be any term.

The invertibility of R3 is immediate from (10.4) by taking v* as t.

(10.4) is most conveniently proven by first establishing the following as a lemma.

(10.5) If there is a closed tree for P, then there is a closed tree for P(v//t) where v is any variable and t is any term.

Let T be a closed tree for P. Let us assume that no

variable introduced into a formula of T by R3 occurs in t. As we claimed in footnote 2 of this section, it is always possible to systematically change such variables so that they are all distinct from variables which occur in t.

We can prove (10.5) by induction on the weight of T. Suppose that T has a weight of one. If P contains a disjunct of the form $\ulcorner \text{'R'} \neq \text{'S'} \urcorner$, then P(v//t) contains the same disjunct. If P contains $\ulcorner u = u \urcorner$ as a disjunct, then P(v//t) contains $\ulcorner u(v//t) = u(v//t) \urcorner$ as a disjunct. If P contains disjuncts of the form Q and $\ulcorner -Q \urcorner$, then P(v//t) contains Q(v//t) and $\ulcorner -Q(v//t) \urcorner$ as disjuncts. In each case where P contains disjunct(s) which make P a tree axiom, P(v//t) contains disjunct(s) which make P(v//t) a tree axiom.

We assume that (10.5) holds for every formula which has a closed tree of weight less than n. Let T be a closed tree for P of weight n.

Inspection of the forms of the rules shows that if the first rule used in T is any rule other than R2, the successor of P in T may be written as P* where P*(v//t) and P(v//t) are also related as premise to conclusion of one of our rules. (For this latter claim to hold when P* results from P by R3 it is necessary that the variable introduced by R3 not occur in t.) By the induction hypothesis there is a closed tree for P*(v//t). To the base of

this latter tree we may subjoin an instance of P(v//t).

The argument is similar if the first rule used in T is R2, the only difference being that P has two premises.

Now that (10.5) is established we may proceed to the proof of (10.4).

Let T be a closed tree for $\ulcorner\neg(Ev)P \vee Q\urcorner$. Our proof is by induction on the weight of T. Suppose that T has a weight of one. Then either Q contains disjunct(s) which make Q a tree axiom, in which case there is obviously a closed tree for $\ulcorner\neg P(v//t) \vee Q\urcorner$, or else Q contains either $\ulcorner(Ev)P\urcorner$ or $\ulcorner\neg\neg(Ev)P\urcorner$ as a disjunct.

If Q contains $\ulcorner(Ev)P\urcorner$ as a disjunct, then Q may be written as $\ulcorner(Ev)P \vee R\urcorner$ for some (possibly void) R. We may apply R4 to $\ulcorner\neg P(v//t) \vee (Ev)P \vee R\urcorner$ to obtain $\ulcorner\neg P(v//t) \vee (Ev)P \vee P(v//t) \vee R\urcorner$, the result being a closed tree for $\ulcorner\neg P(v//t) \vee Q\urcorner$. If Q contains $\ulcorner\neg\neg(Ev)P\urcorner$ as a disjunct we first apply R1 and then R4.

We assume that (10.4) holds whenever there is a closed tree for a formula of the form $\ulcorner\neg(Ev)P \vee Q\urcorner$ of weight less than n.

Let T be a closed tree for $\ulcorner\neg(Ev)P \vee Q\urcorner$ of weight n. Inspection of the forms of the rules shows that if the first rule used in T is any rule except R2, R3 where $\ulcorner\neg(Ev)P\urcorner$ is operated upon, or a substitution rule where a

substitution is made in $\ulcorner\neg(Ev)P\urcorner$; then the successor of $\ulcorner\neg(Ev)P \vee Q\urcorner$ in T may be written as $\ulcorner\neg(Ev)P \vee Q^*\urcorner$ where $\ulcorner\neg P(v//t) \vee Q^*\urcorner$ and $\ulcorner\neg P(v//t) \vee Q\urcorner$ are also related as premise to conclusion of one of our rules.

Removing $\ulcorner\neg(Ev)P \vee Q\urcorner$ from the base of T leaves us with a closed tree for $\ulcorner\neg(Ev)P \vee Q^*\urcorner$ of weight n-1. By the induction hypothesis there is a closed tree for $\ulcorner\neg P(v//t) \vee Q^*\urcorner$. To the base of this latter tree we may subjoin an instance of $\ulcorner\neg P(v//t) \vee Q\urcorner$.

If the first rule used in T is R2 the argument is similar, except that $\ulcorner\neg(Ev)P \vee Q\urcorner$ has two premises.

If the first rule used in T is a substitution rule and a substitution is made in $\ulcorner\neg(Ev)P\urcorner$, then the successor of $\ulcorner\neg(Ev)P \vee Q\urcorner$ in T may be written as $\ulcorner\neg(Ev)P^* \vee Q\urcorner$ where $\ulcorner\neg P^*(v//t) \vee Q\urcorner$ and $\ulcorner\neg P(v//t) \vee Q\urcorner$ are also related as premise and conclusion of one of our rules.

Removing $\ulcorner\neg(Ev)P \vee Q\urcorner$ from the base of T leaves us with a closed tree for $\ulcorner\neg(Ev)P^* \vee Q\urcorner$ of weight n-1. By the induction hypothesis there is a closed tree for $\ulcorner\neg P^*(v//t) \vee Q\urcorner$. To the base of this latter tree we may subjoin an instance of $\ulcorner\neg P(v//t) \vee Q\urcorner$.

If the first rule used in T is R3 and $\ulcorner\neg(Ev)P\urcorner$ is operated upon, then the successor of $\ulcorner\neg(Ev)P \vee Q\urcorner$ in T is $\ulcorner\neg P(v//v^*) \vee Q\urcorner$. Removing $\ulcorner\neg(Ev)P \vee Q\urcorner$ from the base of T

leaves us with a closed tree for $\ulcorner$-P(v//v*) v Q$\urcorner$. By (10.5) there is a closed tree for $\ulcorner$(-P(v//v*) v Q)(v*//t)$\urcorner$. But since v* has no free occurrences in either P or Q, $\ulcorner$(-P(v//v*) v Q)(v*//t)$\urcorner$ is $\ulcorner$-P(v//t) v Q$\urcorner$. This completes our proof of (10.4).

To complete our proof of (10.3) we must show the invertibility of R7 and R8.

It is clear that R7 is invertible when P has a closed tree since $\ulcorner$a = a v Q$\urcorner$ is a tree axiom. For the same reason R8 is invertible when P does not have a closed tree. The difficult cases are to show that R7 is invertible when P does not have a closed tree and that R8 is invertible when P does have a closed tree. In these two cases the invertibility of R7 and R8 will be corollaries to the following two results.

(10.6) If there is no closed tree for P but there is a closed tree for $\ulcorner$N('P') v Q$\urcorner$, then there is a closed tree for Q.

(10.7) If there are closed trees for both P and $\ulcorner$-N('P') v Q$\urcorner$, then there is a closed tree for Q.

Unfortunately we can not prove (10.6) and (10.7) by using the simple type of induction upon the weight of a tree which we employed earlier. The first rule used in the closed tree for $\ulcorner$N('P') v Q$\urcorner$, for example, may be R5

where the successor of $\ulcorner N('P') \vee Q \urcorner$ is $\ulcorner N(t) \vee Q \urcorner$ and t is not $\ulcorner 'R' \urcorner$ for any R. The proofs of (10.6) and (10.7) will have to employ a different technique. Fortunately the proofs are sufficiently similar that we need consider only (10.6).

Let T be a closed tree for $\ulcorner N('P') \vee Q \urcorner$. Let us call a tree explicit if the substitutions permitted by R9, R10, R11, and R12 are made only in disjuncts of the form $\ulcorner t \neq t \urcorner$.

We now claim that every formula which has a closed tree also has a closed explicit tree. For suppose that in the original tree a rule such as R9 was used to obtain $\ulcorner P(neg('R')/'-R') \urcorner$ from P. We may eliminate this use of R9 by using R16 to obtain $\ulcorner P \vee neg('R') \neq neg('R') \urcorner$ from P, then using R9 to obtain $\ulcorner P \vee neg('R') \neq '-R' \urcorner$, and finally using R5 to obtain $\ulcorner P(neg('R')/'-R') \vee neg('R') \neq '-R' \urcorner$. For convenience we will suppose that T is an explicit tree.

Let us call a tree frugal if every inequality which is used to make a substitution by R5 is "saved" in the sense of appearing as a disjunct in each later formula in the branch of the formula in which the R5 substitution was made. Not every closed tree is frugal. $\ulcorner t \neq u \urcorner$ may be used at a certain stage to make a substitution by R5, and then a substitution may subsequently be made in

$\ulcorner t \neq u \urcorner$ itself.

We claim, however, that every closed tree can easily be converted into a closed frugal tree. After each use of R5 in the original tree we simply insert a use of R13 in order to get a "copy" of the inequality used to make the R5 substitution. Even if the original inequality is subsequently altered, the copy will be preserved.

It is clear that if the above procedure is applied to an explicit closed tree the result will be a closed tree which is both explicit and frugal. We may assume, then, that T is both explicit and frugal.

In what follows our discussion will be somewhat informal but, we hope, sufficiently clear. By the <u>history</u> of the disjunct $\ulcorner N('P') \urcorner$ in a branch of T we mean the disjuncts for which $\ulcorner N('P') \urcorner$ is "responsible" in that branch. We imagine these disjuncts to be arranged in ascending order.

Suppose that there are m levels in a particular branch of T. Then the history of $\ulcorner N('P') \urcorner$ in that branch will have one of the following three forms.

(10.8) $N(t_1), N(t_2), \ldots, N(t_m)$

(10.9) $N(t_1), \ldots, N(t_{m-1}), a = a$

or

(10.10) $N(t_1)$, ..., $N(t_k)$, $a \neq a$, ..., $a \neq a$.

In each of these types of history t_1 is $\ulcorner$'P'$\urcorner$.

Our method of proof consists of removing from the formulas of T all of the disjuncts of the form $\ulcorner N(t_i)\urcorner$ for which $\ulcorner N(\text{'P'})\urcorner$ was responsible and then showing that this altered tree can be extended into a closed tree for Q.

Suppose that all such disjuncts of the form $\ulcorner N(t_i)\urcorner$ have been removed from T. It seems clear that, with the exception of the terminal formulas in branches which had histories of the form of (10.9), we have a "beginning" tree for Q. That is, with the exception just noted, each formula of the altered tree and its predecessor stand in the relation of premise to conclusion of one of our rules.

In branches which had histories of the form of (10.10) the disjuncts "$a \neq a$" can always be regarded as having been introduced by R16. The terminal formulas of branches of the altered tree which had histories of this form must be tree axioms, since the terminal formulas of the altered branches are the same as the terminal formulas of the original closed tree.

We now consider the branches of the altered tree which had histories of the form of (10.9). We wish to

show that we may extend such branches so that they terminate in tree axioms. In a branch which had such a history t_{m-1} must have been $\ulcorner$'R'$\urcorner$ for some R which had a closed tree. Since P was assumed not to have had a closed tree, P and R must be different sentences.

We assumed that T was explicit, which means that R must have been "derived" from P by some sequence of substitutions each of which was justified by R5. Since we also assumed that T was frugal, each of the inequalities used to effect this derivation is a disjunct of the terminal formula of the altered branch.

Our first step in extending the altered branch which had a history of the form of (10.9) is to replace the disjunct "a = a" in the terminal formula by $\ulcorner$'P' $\neq$ 'P'$\urcorner$. The terminal formula of the altered branch is now related to its predecessor as premise to conclusion of R16. We then make the substitutions needed to derive $\ulcorner$'R'$\urcorner$ from $\ulcorner$'P'$\urcorner$ in the right-hand occurrence of $\ulcorner$'P'$\urcorner$ in $\ulcorner$'P' $\neq$ 'P'$\urcorner$. This procedure will result in the altered branch being extended to a formula which contains $\ulcorner$'P' $\neq$ 'R'$\urcorner$ as a disjunct.

Finally we consider the branches of the altered tree which had histories of the form of (10.8). By the conditions for tree axiomhood it is clear that the only way that the original terminal formula could have been a tree

axiom without the altered terminal formula also being a tree axiom is if the terminal formula of the altered branch contains $\ulcorner\text{-}N(t_m)\urcorner$ as a disjunct.

Here again t_m is derivable from $\ulcorner\text{'}P\text{'}\urcorner$ by a sequence of R5 substitutions using inequalities which appear as disjuncts of the terminal formula. By repeated uses of R6 we may commute each of these inequalities. Since a term of the form (t(u/w))(w/u) may always be t, we may make the "reverse" R5 substitutions in $\ulcorner\text{-}N(t_m)\urcorner$ until we obtain a formula with $\ulcorner\text{-}N(t_1)\urcorner$, that is, with $\ulcorner\text{-}N(\text{'}P\text{'})\urcorner$ as a disjunct. It was assumed that P had no closed tree. This means that we may use R8 to obtain a formula with "a = a" as a disjunct.

This completes our proof of (10.6) which, in turn, completes our proof of (10.3).

We are finally in a position to begin our proof of (10.1), the claim that every formula of SM is eliminable. We will prove (10.1) by performing a "double" induction. We want to show that for every k and for every n, if P is of degree k, then P is n-eliminable. Accordingly, we shall establish the following result.

(10.11) If k and n are integers such that (1) every formula of degree less than k is eliminable and (2) every formula of degree k is (n-1)-eliminable, then every formula of degree k is n-eliminable.

We assume that the antecedent conditions of (10.11) are met, that P is of degree k, that T_1 is a closed tree for $\ulcorner P \vee Q \urcorner$ of weight n, and that T_2 is a closed tree for $\ulcorner \neg P \vee Q \urcorner$.

The main division in our proof of (10.11) is in first considering the case where k is zero and then considering the case where k is greater than zero.

Suppose that k is zero. Then P is either of the form $\ulcorner t = u \urcorner$ or of the form $\ulcorner N(t) \urcorner$.

Suppose that n is one and that P is of the form $\ulcorner t = u \urcorner$. By the conditions for tree axiomhood there are three possibilities. Q may contain disjunct(s) which make Q a tree axiom, in which case there is obviously a closed tree for Q; Q may contain $\ulcorner t \neq u \urcorner$ as a disjunct; or t and u may be identical.

Suppose that Q contains $\ulcorner t \neq u \urcorner$ as a disjunct. Then Q may be written as $\ulcorner t \neq u \vee R \urcorner$ for some (possibly void) R. T_2 is a closed tree for $\ulcorner t \neq u \vee Q \urcorner$, that is, for $\ulcorner t \neq u \vee t \neq u \vee R \urcorner$. R13 allows us to subjoin $\ulcorner t \neq u \vee R \urcorner$ to the base of T_2, the result being a closed tree for Q.

Suppose that t and u are identical. Then T_2 is a closed tree for $\ulcorner t \neq t \vee Q \urcorner$. R16 allows us to subjoin an instance of Q to T_2.

Suppose that n is one and that P is of the form $\ulcorner N(t) \urcorner$. Here there are only two possibilities. Either Q contains disjunct(s) which make Q a tree axiom, the trivial case, or else Q contains $\ulcorner \neg N(t) \urcorner$ as a disjunct.

Suppose that Q does contain $\ulcorner \neg N(t) \urcorner$ as a disjunct. Then Q may be written as $\ulcorner \neg N(t) \vee R \urcorner$ for some (possibly void) R. T_2 is a closed tree for $\ulcorner \neg N(t) \vee Q \urcorner$, that is, for $\ulcorner \neg N(t) \vee \neg N(t) \vee R \urcorner$. R14 allows us to subjoin $\ulcorner \neg N(t) \vee R \urcorner$ to the base of T_2, the result being a closed tree for Q.

We now suppose that n is greater than one. We consider the various possibilities for the first rule used in T_1. Inspection of the forms of the rules shows that if the first rule used in T_1 is any rule other than R2, a substitution rule where a substitution is made in P, or a rule whereby P is operated upon; then the successor of $\ulcorner P \vee Q \urcorner$ in T_1 may be written as $\ulcorner P \vee Q^* \urcorner$ where Q^* and Q, as well as $\ulcorner \neg P \vee Q^* \urcorner$ and $\ulcorner \neg P \vee Q \urcorner$, are also related as premise to conclusion of one of our rules.

Removing $\ulcorner P \vee Q \urcorner$ from the base of T_1 leaves us with a closed tree for $\ulcorner P \vee Q^* \urcorner$ of weight n-1. T_2 is a closed tree for $\ulcorner \neg P \vee Q \urcorner$. By the Invertibility Principle there is a closed tree for $\ulcorner \neg P \vee Q^* \urcorner$. By antecedent condition (2) of (10.11) P is (n-1)-eliminable, which means that there is a closed tree for Q^*. To the base of this lat-

ter tree we may subjoin an instance of Q.

If the first rule used in T_1 is R2 the argument is similar, except that there are two premises for $\ulcorner P \vee Q \urcorner$.

If the first rule used in T_1 is a substitution rule where a substitution is made in P, then the successor of $\ulcorner P \vee Q \urcorner$ in T_1 may be written as $\ulcorner P^* \vee Q \urcorner$ where $\ulcorner -P^* \vee Q \urcorner$ and $\ulcorner -P \vee Q \urcorner$ are also related as premise to conclusion of one of our rules, and the degree of P^* is zero.

Removing $\ulcorner P \vee Q \urcorner$ from the base of T_1 leaves us with a closed tree for $\ulcorner P^* \vee Q \urcorner$ of weight n-1. T_2 is a closed tree for $\ulcorner -P \vee Q \urcorner$. By the Invertibility Principle there is a closed tree for $\ulcorner -P^* \vee Q \urcorner$. By antecedent condition (2) of (10.11) P^* is (n-1)-eliminable, which means that there is a closed tree for Q.

Since P is either of the form $\ulcorner t = u \urcorner$ or of the form $\ulcorner N(t) \urcorner$, the only way in which P can be operated upon is for P to be $\ulcorner N('R') \urcorner$ for some sentence R and for the first rule to be used in T_1 to be R7.

If R has no closed tree, then the successor of $\ulcorner P \vee Q \urcorner$ in T_1 is $\ulcorner a \neq a \vee Q \urcorner$. Removing $\ulcorner P \vee Q \urcorner$ from the base of T_1 leaves us with a closed tree for $\ulcorner a \neq a \vee Q \urcorner$. R16 allows us to subjoin an instance of Q to the base of this latter tree.

If R does have a closed tree then appeal to (10.7)

and the fact that T_2 is a closed tree for $\ulcorner$-N('R') v Q$\urcorner$ shows that there is a closed tree for Q.

This completes our proof of (10.11) for the case in which k, the degree of P, is zero. We now consider the case in which k is greater than zero.

Let P be of degree k for some k greater than zero. Then P must be of one of three forms: $\ulcorner$-R$\urcorner$, $\ulcorner$R v S$\urcorner$, or $\ulcorner$(Ev)R$\urcorner$. We shall consider each of these cases in turn.

Suppose that P is $\ulcorner$-R$\urcorner$ for some R. Then T_1 is a closed tree for $\ulcorner$- -R v Q$\urcorner$. By the Invertibility Principle there is a closed tree for $\ulcorner$R v Q$\urcorner$, since $\ulcorner$R v Q$\urcorner$ is related to $\ulcorner$- -R v Q$\urcorner$ as premise to conclusion of R1. The degree of R is less than k. By antecedent condition (1) of (10.11) R is eliminable, which means that there is a closed tree for Q.

Suppose that P is $\ulcorner$R v S$\urcorner$ for some R and some S. Then T_1 is a closed tree for $\ulcorner$R v S v Q$\urcorner$ and T_2 is a closed tree for $\ulcorner$-(R v S) v Q$\urcorner$. By the Invertibility Principle there is a closed tree for $\ulcorner$-R v Q$\urcorner$, since $\ulcorner$-R v Q$\urcorner$ is related to $\ulcorner$-(R v S) v Q$\urcorner$ as premise to conclusion of R2. By the Extra Disjunct Principle there is a closed tree for $\ulcorner$-R v S v Q$\urcorner$. The degree of R is less than k. By antecedent condition (1) of (10.11) R is eliminable, which means that there is a closed tree for $\ulcorner$S v Q$\urcorner$.

We appeal once more to the Invertibility Principle. ⌜-S v Q⌝ is also related to ⌜-(R v S) v Q⌝ as premise to conclusion of R2, which means that there is a closed tree for ⌜-S v Q⌝. The degree of S is less than k. By antecedent condition (1) of (10.11) S is eliminable, which means that there is a closed tree for Q.

Finally we suppose that P is ⌜(Ev)R⌝ for some v and some R. Suppose that n, the weight of T_1, is one. ⌜(Ev)R v Q⌝ is a tree axiom only if Q contains disjunct(s) which make Q a tree axiom, in which case there is obviously a closed tree for Q, or if Q contains ⌜-(Ev)R⌝ as a disjunct.

Suppose that Q does contain ⌜-(Ev)R⌝ as a disjunct. Then Q may be written as ⌜-(Ev)R v S⌝ for some (possibly void) S. T_2 is a closed tree for ⌜-(Ev)R v -(Ev)R v S⌝. By R15 we may subjoin an instance of ⌜-(Ev)R v S⌝ to the base of T_2, the result being a closed tree for Q.

Suppose that the weight of T_1 is greater than one. Inspection of the forms of the rules shows that if the first rule used in T_1 is any rule other than R2, a substitution rule by which a substitution is made in R, or R4 where ⌜(Ev)R⌝ is operated upon; then the successor of ⌜(Ev)R v Q⌝ in T_1 may be written as ⌜(Ev)R v Q*⌝ where Q* is related to Q, and ⌜-(Ev)R v Q*⌝ is related to ⌜-(Ev)R v Q⌝, as premise to conclusion of one of our

rules.

Removing $\ulcorner(Ev)R \vee Q\urcorner$ from the base of T_1 leaves us with a closed tree for $\ulcorner(Ev)R \vee Q^*\urcorner$ of weight n-1. T_2 is a closed tree for $\ulcorner\neg(Ev)R \vee Q\urcorner$. By the Invertibility Principle there is a closed tree for $\ulcorner\neg(Ev)R \vee Q^*\urcorner$. By antecedent condition (2) of (10.11) $\ulcorner(Ev)R\urcorner$ is (n-1)-eliminable, which means that there is a closed tree for Q*. To the base of this latter tree we may subjoin an instance of Q.

If the first rule used in T_1 is R2 the argument is similar, except that $\ulcorner(Ev)R \vee Q\urcorner$ has two premises.

If the first rule used in T_1 is a substitution rule by which a substitution is made in R, then the successor of $\ulcorner(Ev)R \vee Q\urcorner$ in T_1 may be written as $\ulcorner(Ev)R^* \vee Q\urcorner$ where the degree of $\ulcorner(Ev)R^*\urcorner$ is k and $\ulcorner\neg(Ev)R^* \vee Q\urcorner$ is related to $\ulcorner\neg(Ev)R \vee Q\urcorner$ as premise to conclusion of one of our rules.

Removing $\ulcorner(Ev)R \vee Q\urcorner$ from the base of T_1 leaves us with a closed tree for $\ulcorner(Ev)R^* \vee Q\urcorner$ of weight n-1. T_2 is a closed tree for $\ulcorner\neg(Ev)R \vee Q\urcorner$. By the Invertibility Principle there is a closed tree for $\ulcorner\neg(Ev)R^* \vee Q\urcorner$. By antecedent condition (2) of (10.11) $\ulcorner(Ev)R^*\urcorner$ is (n-1)-eliminable, which means that there is a closed tree for Q.

Suppose that the first rule used in T_1 is R4 and that $\ulcorner(Ev)R\urcorner$ is operated upon. Then the successor of $\ulcorner(Ev)R \vee Q\urcorner$ in T_1 may be written as $\ulcorner(Ev)R \vee R(v//t) \vee Q\urcorner$. Removing $\ulcorner(Ev)R \vee Q\urcorner$ from the base of T_1 leaves us with a closed tree for $\ulcorner(Ev)R \vee R(v//t) \vee Q\urcorner$ of weight n-1.

T_2 is a closed tree for $\ulcorner\neg(Ev)R \vee Q\urcorner$. By the Extra Disjunct Principle there is a closed tree for $\ulcorner\neg(Ev)R \vee R(v//t) \vee Q\urcorner$. By antecedent condition (2) of (10.11) $\ulcorner(Ev)R\urcorner$ is (n-1)-eliminable, which means that there is a closed tree for $\ulcorner R(v//t) \vee Q\urcorner$. An appeal to (10.4) and the fact that T_2 is a closed tree for $\ulcorner\neg(Ev)R \vee Q\urcorner$ shows that there is a closed tree for $\ulcorner\neg R(v//t) \vee Q\urcorner$. By antecedent condition (1) of (10.11) R(v//t) is eliminable, since the degree of R(v//t) is k-1, which means that there is a closed tree for Q.

This completes our proof of (10.11). (10.1), our original statement of the Cut-Elimination Theorem, is an immediate consequence of (10.11). The proof is by contradiction.

Suppose that not every formula of SM is eliminable. Then there must be an integer k such that there are formulas of degree k which are not eliminable but no formulas of degree less than k which are not eliminable. Let K be the set of formulas of SM of degree k which are

not eliminable. There must be an integer n such that there are members of K which are not n-eliminable, but no members of K which are not (n-1)-eliminable. But this contradicts (10.11) which says that for such a k and such an n, every formula of degree k is n-eliminable.

/11/

THE I-EXTENSIONS OF SM

The consistency of SM, which was proven in the preceding section, is sufficient to refute (6.3), the strong form of the Parity of Naming Thesis. Refuting (6.3), however, was not our sole purpose in constructing the theory SM. We also want to use SM to answer our original question from Section One about the existence of a modal antinomy.

In Section One we asked whether a contradiction could be derived from the identity statements resulting from the placement of modal sentences on a certain blackboard, together with instances of the schemata from a standard theory of modal logic.

SM contains as theorems the syntactical analogues of the instances of all of the modal schemata of Lewis' system S5. In order to represent the identity statements resulting from the placement of sentences on the blackboard we introduce the idea of an I-extension of SM.

A theory will be said to be an I-extension of SM if its non-logical axioms are the non-logical axioms of SM together with a finite set of axioms, to be called I-axioms, of the form $\ulcorner a_i = \text{'}P_i\text{'} \urcorner$ where a_i is a constant and P_i is a sentence of SM. It is also stipulated that

no constant may have more than one free occurrence in the set of I-axioms.

The non-logical inference (7.7) must also be altered for the I-extensions of SM. For any proof in an I-extension of SM we may define certain lines of the proof to be I-dependent. A line is to be I-dependent if it is either an I-axiom or else follows by modus ponens or universal generalization from at least one earlier line of the proof which is I-dependent.

For the I-extensions of SM we substitute the following non-logical inference for (7.7).

(11.1) From P to infer $\ulcorner N('P') \urcorner$, if P is a sentence which is not I-dependent.

The effect of (11.1) is to allow us to apply our necessitation rule only to sentences which are theorems of SM. The purpose of this restriction is to prevent us from making inferences in our formal theory which would correspond to the inference from

(11.2) The first sentence on B is "The first sentence on B is necessary"

to

(11.3) "The first sentence on B is 'The first sentence on B is necessary'" is necessary.

This inference is intuitively undesirable as there seems to be no necessity in the way that the chalk is arranged on some particular blackboard.

In Section One we offered the following arrangement of sentences as a possible candidate for generating a modal antinomy.

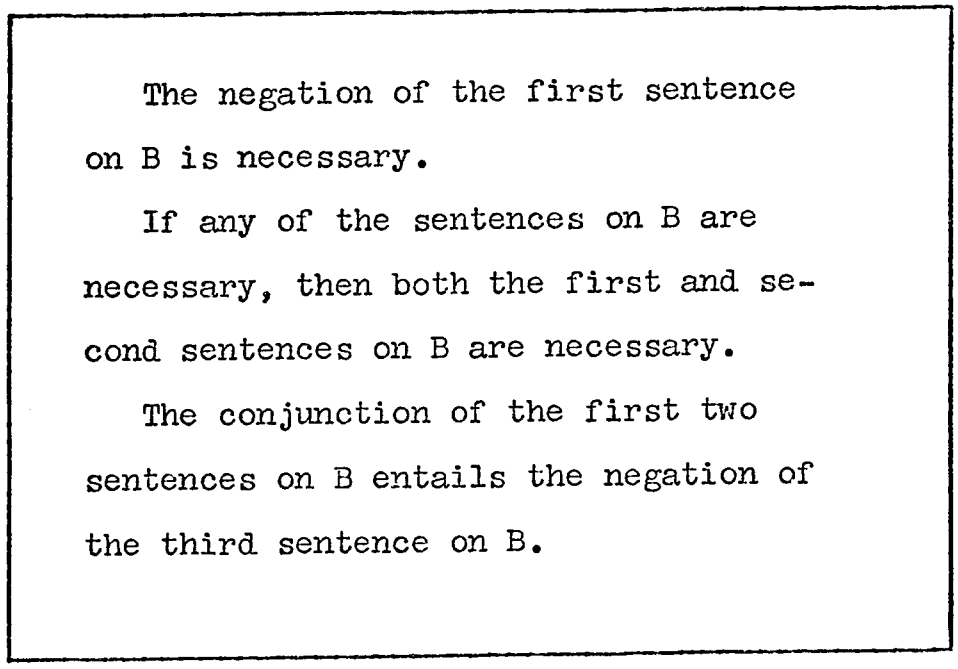


Figure 1

Letting the intended interpretations of "a", "b", and "c" be the first, second, and third sentences respectively on B, we can see that the possibility of generating a modal antinomy from the placement of the sentences in Figure 1 reduces to the question of the consistency of the I-extension of SM whose I-axioms are the following.

(11.4) a = 'N(neg(a))'

b = '(N(a) v N(b) v N(c)) $\longrightarrow$ (N(a) & N(b))'

c = 'N(cond(conj(a,b),neg(c)))'.

The more general question is "Are all of the I-extensions of SM consistent?" It seems reasonable to regard an affirmative answer to this question as equivalent to a negative answer to the question "Is there a modal antinomy?" In the next section we shall prove that every I-extension of SM is, in fact, consistent.

We shall conclude this section by contrasting SM with another theory in L, ST, intended as a formalization of a syntactical theory of truth. ST is like SM except for containing the one-place predicate symbol "T" instead of "N". ST lacks the non-logical inference (7.7) and has instead of axioms (1) through (3), the modal schemata, of (7.6) the single schema

(11.5) P $\longleftrightarrow$ T('P').

To formally construct the antinomy of the liar we consider the I-extension of ST whose only I-axiom is "a = '-T(a)''. A contradiction is easily derived in this extension.

(11.6)	1.	-T(a) $\longleftrightarrow$ T('-T(a)')	axiom
	2.	a = '-T(a)'	I-axiom
	3.	-T(a) $\longleftrightarrow$ T(a)	1,2, identity

There is a classical antinomy involving indirect self reference in which a first sentence asserts the truth of a second while the second denies the truth of the first. To formally construct this antinomy we consider the I-extension of ST whose I-axioms are "a = 'T(b)'" and "b = '-T(a)".

(11.7)	1.	-T(a) $\leftrightarrow$ T('-T(a)')	axiom
	2.	b = '-T(a)'	I-axiom
	3.	-T(a) $\leftrightarrow$ T(b)	1,2, identity
	4.	T(b) $\leftrightarrow$ T('T(b)')	axiom
	5.	a = 'T(b)'	I-axiom
	6.	T(b) $\leftrightarrow$ T(a)	4,5, identity
	7.	-T(a) $\leftrightarrow$ T(a)	3,6

The purpose of our digression into ST was to impress the reader with the plausibility of the supposition that if there were a self referential antinomy at all analogous to an antinomy of truth, it could be represented in an I-extension of SM.

In Section Five we saw that Kaplan and Montague claimed to know of no other way of treating self referential sentences than to be able to construct for any predicate F a sentence S such that S is provably equivalent to F($\overline{S}$), $\overline{S}$ here being the "standard name" of S. It should be noted that this can always be done for any predicate of SM.

Consider the I-extension of SM whose only I-axiom is $\ulcorner a = \text{'F(a)'} \urcorner$. In this I-extension we can prove

(11.8) $F(a) \longleftrightarrow F(\text{'F(a)'})$.

Letting S be $\ulcorner F(a) \urcorner$ we have a sentence S such that S is provably equivalent to $F(\overline{S})$. Thus, SM seems to meet the conditions for expressibility of self referential sentences which were demanded by Kaplan and Montague in Section Five.

A PROOF THAT EVERY I-EXTENSION OF SM IS CONSISTENT

In this section we shall prove that every I-extension of SM is consistent. We consider the I-extension whose I-axioms are $\ulcorner a_1 = \text{'}P_1\text{'}\urcorner$, ..., $\ulcorner a_n = \text{'}P_n\text{'}\urcorner$. We refer to this I-extension as I#.

We let I* be the disjunction $\ulcorner a_1 \neq \text{'}P_1\text{'} \vee \ldots a_n \neq \text{'}P_n\text{'}\urcorner$. We begin our proof by showing that if a formula Q is provable in I#, then there is a closed tree for $\ulcorner \text{I*} \vee Q\urcorner$. We complete the proof by showing that there is at least one formula R such that there is no closed tree for $\ulcorner \text{I*} \vee R\urcorner$.

We first prove

(12.1) If Q is provable in I#, then there is a closed tree for $\ulcorner \text{I*} \vee Q\urcorner$.

Our proof of (12.1) is by induction on the length of the proof for Q. Suppose that Q has a proof whose length is one. Then Q is either an axiom of SM or an I-axiom. If Q is an axiom of SM, then Q has a closed tree and, by the Extra Disjunct Principle, so does $\ulcorner \text{I*} \vee Q\urcorner$. If Q is an I-axiom, then $\ulcorner \text{-}Q\urcorner$ appears as a disjunct in I*, which means that $\ulcorner \text{I*} \vee Q\urcorner$ is a tree axiom.

Suppose that (12.1) holds for every formula which has

a proof of length less than n. Let Q have a proof of exactly length n. If Q is either an axiom of SM or an I-axiom, then there is a closed tree for $\ulcorner$I* v Q$\urcorner$ by the argument of the preceding paragraph.

Suppose that Q follows from an earlier line by our modified non-logical inference (11.1). Then Q is $\ulcorner$N('R')$\urcorner$ for some R which is not I-dependent. Since R is not I-dependent it is a theorem of SM and, hence, has a closed tree. Applying R7 to $\ulcorner$I* v N('R')$\urcorner$ gives us $\ulcorner$I* v a = a$\urcorner$, a tree axiom.

Suppose that Q follows from an earlier line by universal generalization. Then Q is $\ulcorner$-(Ev)-R$\urcorner$ for some R. By the induction hypothesis there is a closed tree for $\ulcorner$I* v R$\urcorner$. R can be represented as R(v//v*) where v* is just v. To the base of the closed tree for $\ulcorner$I* v R(v//v*)$\urcorner$ we first subjoin, by R1, an instance of $\ulcorner$I* v - -R(v//v*)$\urcorner$ and then, by R3, an instance of $\ulcorner$I* v -(Ev)-R$\urcorner$. This last step is legitimate because we know that v has no free occurrences in $\ulcorner$I* v -(Ev)-R$\urcorner$.

Suppose that Q follows from two earlier lines R and $\ulcorner$-R v Q$\urcorner$ by *modus ponens*. By the induction hypothesis there are closed trees for $\ulcorner$I* v -R v Q$\urcorner$ and $\ulcorner$I* v R$\urcorner$. By the Extra Disjunct Principle there is a closed tree for $\ulcorner$I* v R v Q$\urcorner$. By the Cut-Elimination Theorem there is a closed tree for $\ulcorner$I* v Q$\urcorner$.

This completes our proof of (12.1).

Let us say that a formula of the form $\ulcorner a_1 \neq \text{'}P_1\text{'} \vee \ldots \vee a_m \neq \text{'}P_m\text{'} \vee a_{m+1} \neq a_{m+1}\urcorner$, where each a_i, $1 \leq i \leq m+1$, is a distinct individual constant of SM, is amiable. We now prove

(12.2) If P is amiable, then there is no closed tree for P.

It will follow from (12.2) that there is no closed tree for $\ulcorner I^* \vee a_{n+1} \neq a_{n+1}\urcorner$, which means that $\ulcorner a_{n+1} \neq a_{n+1}\urcorner$ is not a theorem of I# and that I# is consistent.

Suppose that there was a closed tree for some amiable formula. We would like to make some comments on the nature of this closed tree.

The only rules that could have been used in such a closed tree are R5, R6, R9, R10, R11, R12, R13, and R16. The only disjuncts which could appear in the formulas of such a tree are disjuncts of the form $\ulcorner t \neq u\urcorner$. Such a tree could be closed only by virtue of having as a disjunct of its terminal formula a disjunct of the form $\ulcorner \text{'}Q\text{'} \neq \text{'}R\text{'}\urcorner$ where Q and R are different sentences.

We also note that in such a closed tree there would be no need to introduce a disjunct of the form $\ulcorner a_i \neq a_i\urcorner$ or $\ulcorner \text{'}P_i\text{'} \neq \text{'}P_i\text{'}\urcorner$ by R16. If such a disjunct was essen-

tial to the tree it could be obtained from $\ulcorner a_i \neq \text{'}P_i\text{'} \urcorner$ by using R13, perhaps R6, and R5 in an obvious way. In the supposed closed trees for amiable formulas which we consider we will suppose that no disjunct introduced by R16 is $\ulcorner a_i \neq a_i \urcorner$ or $\ulcorner \text{'}P_i\text{'} \neq \text{'}P_i\text{'} \urcorner$.

In an obvious way we can single out one disjunct as the disjunct from which the disjunct of the form $\ulcorner \text{'}Q\text{'} \neq \text{'}R\text{'} \urcorner$ was originally "derived". We will refer to this disjunct as the tree's fractious disjunct. The fractious disjunct will be of one of two sorts. It will either be $\ulcorner a_i \neq \text{'}P_i\text{'} \urcorner$ or it will be a disjunct of the form $\ulcorner t \neq t \urcorner$ which was introduced by R16. We will refer to disjuncts of these two types as the tree's possibly fractious disjuncts.

We now introduce the notion of the rank of a disjunct for possibly fractious disjuncts. For a disjunct of the form $\ulcorner a_i \neq \text{'}P_i\text{'} \urcorner$, the rank of the disjunct is the number of free occurrences of "-" and "v" in P_i which do not occur in any part of P_i of the form $\ulcorner (Ev)Q \urcorner$. If a possibly fractious disjunct of the form $\ulcorner t \neq t \urcorner$ is $\ulcorner \text{'}R\text{'} \neq \text{'}R\text{'} \urcorner$ for some R, then the rank of the disjunct is again the number of free occurrences of "-" and "v" in R which do not occur in any part of R of the form $\ulcorner (Ev)Q \urcorner$. If $\ulcorner t \neq t \urcorner$ is not of the form $\ulcorner \text{'}R\text{'} \neq \text{'}R\text{'} \urcorner$, then the rank of the disjunct is one less than the number of free occurrences of terms in t. This latter reduction by one is made so that

the lowest rank of either type disjunct is zero.

Let u be one of the terms a_i or t where $\ulcorner a_i \neq {}'P_i{}'\urcorner$ or $\ulcorner t \neq t \urcorner$ is a possibly fractious disjunct. We specify a set of terms A_u as follows.

(12.3) A_u contains all and only those terms so specified by the following eight conditions.

(1) If u is a_i, then both a_i and $\ulcorner {}'P_i{}' \urcorner$ are in A_u.

(2) If u is t, then t is in A_u.

If w is a term in A_u, then all instances of the following are in A_u:

(3) w('Q v R'/disj('Q','R'))

(4) w(disj('Q','R')/'Q v R')

(5) w('-Q'/neg('Q'))

(6) w(neg('Q')/'-Q')

(7) $w(a_j/{}'P_j{}')$, for $1 \leq j \leq m$

(8) $w({}'P_j{}'/a_j)$, for $1 \leq j \leq m$.

Every inequality which is derivable by R1-R16 from a possibly fractious disjunct whose set of associated terms is A_u will be of the form $\ulcorner r \neq s \urcorner$ where both r and s are in A_u. The converse, although not necessary to our proof, is also true; that if r and s are both in A_u, then $\ulcorner r \neq s \urcorner$ is derivable by R1-R16 as a disjunct. Both of these claims can be proven by simple inductions. In the first case the induction is over the length of the deri-

vation by R1-R16. In the second case the induction is over the number of times that the clauses of (12.3) must be applied in order to obtain r and s from u.

The claim that there is a closed tree for some amiable formula can now be reduced to the equivalent claim that there is a set of terms A_u which contains terms of the form $\ulcorner{}'Q'\urcorner$ and $\ulcorner{}'R'\urcorner$, where Q and R are different sentences.

We supposed that there were closed trees for some amiable formulas. Of all these closed trees, there must be at least one whose fractious disjunct is of no greater rank than the fractious disjunct of any other closed tree for an amiable formula. Call this tree T. We first consider the possibly fractious disjuncts of T of rank zero.

Suppose that the rank of a possibly fractious disjunct of T of the form $\ulcorner a_i \neq {}'P_i'\urcorner$ is zero. By the conditions for constructing A_{a_i}, this set can contain only a_i, $\ulcorner{}'P_i'\urcorner$, and those a_j, $1 \leq j \leq m$, such that $\ulcorner a_j \neq {}'P_i'\urcorner$ is also a possibly fractious disjunct of T. Thus A_{a_i} can not contain $\ulcorner{}'Q'\urcorner$ and $\ulcorner{}'R'\urcorner$ for two different sentences Q and R.

Suppose that the rank of a possibly fractious disjunct of T of the form $\ulcorner t \neq t\urcorner$, where t is $\ulcorner{}'R'\urcorner$ for some R, is zero. Since R is known to be distinct from each

P_j, $1 \leq j \leq m$, A_t contains only the term $\ulcorner$'R'$\urcorner$. Thus A_t can not contain $\ulcorner$'Q'$\urcorner$ and $\ulcorner$'R'$\urcorner$ for two different sentences Q and R.

Suppose that the rank of a possibly fractious disjunct of T of the form $\ulcorner t \neq t \urcorner$, where t is not $\ulcorner$'R'$\urcorner$ for any R, is zero. Then t is either a variable or a constant different from each a_j, $1 \leq j \leq m$. In this case A_t contains t as its only term. Here again A_t can not contain $\ulcorner$'Q'$\urcorner$ and $\ulcorner$'R'$\urcorner$ for two different sentences Q and R.

We conclude that T does not have a fractious disjunct of rank zero.

The fractious disjunct of T must be of rank k, where k is greater than zero. This fractious disjunct will be of one of six forms: $\ulcorner a_i \neq$ 'U v W'$\urcorner$, $\ulcorner$disj(u,w) $\neq$ disj(u,w)$\urcorner$, $\ulcorner$'U v W' $\neq$ 'U v W'$\urcorner$, $\ulcorner a_i \neq$ '-U'$\urcorner$, $\ulcorner$neg(u) $\neq$ neg(u)$\urcorner$, or $\ulcorner$'-U' $\neq$ '-U'$\urcorner$.

We consider two cases, according to whether the fractious disjunct is of one of the first three forms mentioned above or whether it is of one of the last three forms. As will become clear from our method of proof, the proofs for the two cases are sufficiently similar that we need consider only the case where the fractious disjunct is of one of the first three forms.

Letting t be a_i, $\ulcorner$disj(u,w)$\urcorner$, or $\ulcorner$'U v W'$\urcorner$ as the

case may be, we establish the following claim.

(12.4) If w is in A_t, then w is either a_j, $1 \leq j \leq m$, or of the form $\ulcorner \urcorner Q \vee R \urcorner$, or of the form $\ulcorner disj(r,s) \urcorner$.

Inspection of the conditions of (12.3) for constructing A_t is sufficient to establish (12.4).

In order for our disjunct of rank k to be the fractious disjunct of T, A_t must contain both $\ulcorner \urcorner Q \urcorner$ and $\ulcorner \urcorner R \urcorner$ for two different sentences Q and R. In virtue of (12.4) $\ulcorner \urcorner Q \urcorner$ and $\ulcorner \urcorner R \urcorner$ must be terms $\ulcorner \urcorner S \vee X \urcorner$ and $\ulcorner \urcorner Y \vee Z \urcorner$ respectively, where $\ulcorner S \vee X \urcorner$ and $\ulcorner Y \vee Z \urcorner$ are different sentences.

In order for $\ulcorner S \vee X \urcorner$ and $\ulcorner Y \vee Z \urcorner$ to be different sentences, either S and Y must be different sentences or X and Z must be different sentences. Our proof proceeds in the same fashion on either supposition, so we suppose that S and Y are different sentences.

Both $\ulcorner \urcorner S \vee X \urcorner$ and $\ulcorner \urcorner Y \vee Z \urcorner$ are derivable from either a_i, $\ulcorner disj(u,w) \urcorner$, or $\ulcorner \urcorner U \vee W \urcorner$, depending on which type of fractious disjunct we have, by a finite sequence of substitutions permitted by the conditions of (12.3).

The case where $\ulcorner \urcorner S \vee X \urcorner$ and $\ulcorner \urcorner Y \vee Z \urcorner$ are derived from a_i can be ignored, since the only substitution that can be made in a_i is to replace a_i by $\ulcorner \urcorner U \vee W \urcorner$. Thus we

may suppose that both $\ulcorner$'S v X'$\urcorner$ and $\ulcorner$'Y v Z'$\urcorner$ are derivable from either $\ulcorner$disj(u,w)$\urcorner$ or $\ulcorner$'U v W'$\urcorner$.

Any sequences of substitutions which will lead from $\ulcorner$disj(u,w)$\urcorner$ or $\ulcorner$'U v W'$\urcorner$ to $\ulcorner$'S v X'$\urcorner$ and $\ulcorner$'Y v Z'$\urcorner$ are sequences of substitutions which will also lead from u or $\ulcorner$'U'$\urcorner$ to $\ulcorner$'S'$\urcorner$ and $\ulcorner$'Y'$\urcorner$. In general, some of the substitutions in the latter sequence will be vacuous. That $\ulcorner$'S'$\urcorner$ and $\ulcorner$'Y'$\urcorner$ can be so derived from u or $\ulcorner$'U'$\urcorner$ can be proven by induction on the lengths of the sequences of substitutions.

The preceding argument shows that we can construct a closed tree for an amiable formula with a fractious disjunct which is either $\ulcorner a_i \neq$ 'U'$\urcorner$, $\ulcorner u \neq u\urcorner$, or $\ulcorner$'U' $\neq$ 'U'$\urcorner$. In each of these cases the closed tree will have a fractious disjunct whose rank is less than k. But this contradicts our assumption that there were no closed trees for amiable formulas with fractious disjuncts of rank less than k. We conclude that there are no amiable formulas with closed trees.

This completes our proof that every I-extension of SM is consistent.

/13/

WEAKENING THE PARITY OF NAMING THESIS

Before explaining what we intend to do in this section it will be useful to review some of what has been accomplished in the preceding sections.

In Section One we characterized as an Antinomist one who believed the necessarily vague claim that the modal predicates, like the predicate "____ is true", can lead to inconsistency when self reference is involved.

In Section One we also gave a precise description of a state of affairs whose existence we would count as establishing the existence of a modal antinomy. In virtue of the proof in Section Twelve that every I-extension of SM is consistent we may categorically assert that there is no modal antinomy of this description.

What we in effect did in Section One was to replace a vague claim by a precise claim. Our intention, obviously enough, was that an answer to the precise claim should also be a reasonable answer to the vague claim. But what remains to be done is to argue that the precise characterization of a modal antinomy in Section One is in fact a reasonable replacement for the vague claim. We must take seriously the following argument which the Antinomist might use.

Of course, the I-extensions of SM are consistent but that does not conclusively refute the Antinomist's thesis because the I-extensions of SM may be too weak to express all of the assumptions needed to derive a contradiction. The consistency of the I-extensions shows that no contradiction can be derived from the instances of the modal schemata together with the identity statements generated by the relative positions of the sentences on the blackboard. But it may be that a contradiction can be derived from the instances of modal schemata and the identity statements together with other quite reasonable assumptions. If a contradiction could be so derived it would show that at least one of the premises of the argument must be abandoned. The identity statements are beyond reproach. So if the, as yet unspecified, additional assumptions are even the least bit more plausible than the self referential instances of the modal schemata, then surely it is the latter which would have to be abandoned.

As an initial remark upon this line of criticism we may point out that the critic has at least gone beyond the original statement of the Antinomist's position. In that statement it was suggested that the predicate "____ is necessary" might lead to a contradiction in the way that the predicate "____ is true" does. But the antinomy involving "____ is true" requires only the identity statements and the self referential instances of the

schema for the use of "____ is true" . The proof of the consistency of the I-extensions of SM shows that the identity statements and the self referential instances of the schemata for the use of "____ is necessary" are not sufficient to construct an antinomy.

The preceding remarks do not show that the Antinomist must be mistaken in thinking that there might be a modal antinomy. They show only that if there is such an antinomy it involves, in an essential way, something more than is involved in the antinomies concerning truth. It is incumbent upon the Antinomist to suggest what this "something more" might be.

The most fruitful place to look for the additional assumptions which might be needed to derive a modal antinomy is the proof of Montague's Theorem. This theorem, which we stated as (4.4), said that formal theories which met certain conditions were inconsistent. The theories shown to be inconsistent were arithmetical theories, with numerals being used as names of expressions.

As we saw in Section Six, Montague claims that similar results can be obtained for theories which are like SM in that they form names of expressions by enclosing those expressions in quotation marks. This claim was referred to as the vague form of the Parity of Naming Thesis.

(6.3) was a proposal of analogous antecedent conditions which a theory $\underline{T}_1$, forming names by means of quotation marks, might have to meet if the results which could be proven about a theory $\underline{T}_2$, meeting the antecedent conditions of Montague's Theorem, were also to be provable for $\underline{T}_1$. We took the strong form of the Parity of Naming Thesis to be the claim that Montague's Theorem implied (6.3). This was equivalent to the claim that (6.3) was true. The proof of the consistency of SM in Section Ten showed that (6.3), and hence the strong form of the Parity of Naming Thesis, was false.

The claim that additional assumptions may be needed to derive a modal antinomy can now be interpreted as the claim that the antecedent conditons of (6.3) are too weak. That is, the antecedent conditions of (6.3) may not be analogous to the antecedent conditions of Montague's Theorem in all relevant respects. To strengthen the antecedent conditions of (6.3) is to weaken the Parity of Naming Thesis since there will be fewer theories which the thesis asserts to be inconsistent.

We must ask how the antecedent conditions of Montague's Theorem might plausibly be claimed to be stronger than the antecedent conditions of (6.3). An obvious answer is that the theories which meet the antecedent conditions of Montague's Theorem must contain arithmetical subtheories, while (6.3) seems to impose no corresponding

requirement.

Suppose that the Parity of Naming Thesis was taken to be the claim that Montague's Theorem implies the following.

(13.1) If $\underline{T}$ is a theory in $\underline{L}$ such that

(1) $\underline{T}$ is an extension of $\underline{Q}$ (or of $\underline{Q}^{(\underline{Z})}$ for some one-place predicate $\underline{Z}$)

and there is a one-place predicate $\underline{N}$ of $\underline{T}$ such that for all sentences P, R of $\underline{T}$

(2) $\vdash_{\underline{T}} \underline{N}(\text{'P'}) \longrightarrow P$

(3) $\vdash_{\underline{T}} \underline{N}(\text{'}\underline{N}(\text{'P'}) \longrightarrow P\text{'})$

(4) $\vdash_{\underline{T}} \underline{N}(\text{'P} \longrightarrow R\text{'}) \longrightarrow (\underline{N}(\text{'P'}) \longrightarrow \underline{N}(\text{'R'}))$

(5) $\vdash_{\underline{T}} \underline{N}(\text{'P'})$, if P is a valid sentence of the first order functional calculus with identity,

then $\underline{T}$ is inconsistent.

(13.1) differs from (6.3) only because of antecedent condition (1) of (13.1), which does not appear as an antecedent condition of (6.3).

We know that $\underline{SM}$ is consistent but yet fulfills the antecedent conditions of (6.3). A natural way to try to show that (13.1) is false is to extend $\underline{SM}$ so that it contains an arithmetic subtheory, and then to try to show that this resulting extension of $\underline{SM}$ is consistent.

Our discussion of the question of the consistency of such an extension of SM will be very informal and very summary. This is done for two reasons. The first reason is that it seems quite clear that such an extension would be consistent, although providing a detailed consistency proof could be a tedious and necessarily non-constructive matter. The second reason is that we can show that the weakening of the Parity of Naming Thesis to (13.1) does not constitute the sort of weakening which would be desired by someone interested in showing that certain syntactical treatments of modality were impossible.

Suppose that SM was extended to a theory SMQ by the addition of a predicate letter "Z" to be interpreted as "____ is a number", and by function letters to be interpreted as the functions for successors, sums, and products. The non-logical axioms of SMQ would be the non-logical axioms of SM plus the non-logical axioms for $Q^{(Z)}$.

We could talk about arbitrary I-extensions of SMQ in the same way that we talked about arbitrary I-extensions of SM. By extending our technique for constructing trees we could show that every I-extension of SMQ was consistent in just the same way that we showed that every I-extension of SM was consistent in Section Twelve.

The additions to the technique for tree construction

would include a rule which allowed us to write $\ulcorner a = a \vee R \urcorner$ as the successor of $\ulcorner P \vee R \urcorner$ if P was a theorem of $\underline{Q}^{(\underline{Z})}$ and to write $\ulcorner a \neq a \vee R \urcorner$ if P was a sentence of $\underline{Q}^{(\underline{Z})}$ but not a theorem of $\underline{Q}^{(\underline{Z})}$. We would also specify that every formula with a disjunct of the form $\ulcorner -\underline{Z}(\text{'P'}) \urcorner$ was a tree axiom. This disjunct would express the intuitive truth that no sentences were numbers.

The reason we can assert with such confidence that every I-extension of SMQ is consistent is that we can see, speaking very loosely again, that the natural "domain of interpretation" for SMQ can be neatly partitioned into two exclusive subdomains: numbers and sentences. That is, we are free to regard every sentence of the form $\ulcorner N(t) \urcorner$ as false where t is an arithmetic term and may regard every sentence of the form $\ulcorner Z(t) \urcorner$ as false where t is not an arithmetic term.

In reading KAPLAN and MONTAGUE (1960) and MONTAGUE (1963) it is clear that the authors' intention in these articles is that the arithmetic subtheory play the role of a *syntax language*. That is, the purpose of $\underline{Q}^{(\underline{Z})}$ is to express certain syntactical properties. Examples of syntactical properties of expressions would be the property of being a variable or the property of being a sentence with exactly two existential quantifiers.

It is true that these syntactical properties can

still be expressed in SMQ via a Godel numbering since every recursive set is representable in SMQ; but the whole point of being able to express these properties is lost. Sentences are the sorts of things of which it makes sense to say that they have certain syntactical properties and of which it also makes sense to say that they are necessary.

In SMQ it is not possible to say that it is the same thing which is both necessary and has a certain syntactical property. SMQ allows us to say that sentences are necessary by predicating "____ is necessary" of the result of enclosing the sentence in quotation marks; but a sentence can be "said" to have one of the other syntactical properties only by predicating an arithmetical predicate of the numeral which is associated with that sentence through a system of Godel numbering.

The charge that the Parity of Naming Thesis is too strong will probably reduce to the charge that a theory such as SM does not contain a sufficient amount of syntax. We could formulate a weaker version of the thesis by replacing condition (1) of (13.1) with the condition (1*): "T contains a reasonable amount of syntax". We will refer to (13.1) with (1*) in place of (1) as the weak form of the Parity of Naming Thesis.

Evaluating the Parity of Naming Thesis now becomes a

more difficult task because we must explicate the notion of "a reasonable amount of syntax". The problems involved in such an explication will be discussed in the next section.

In the remainder of this section we would like to make some comments upon the syntactical strength of SM. In order to assess the criticism that SM might be syntactically deficient we should be aware of just what syntactical notions can be expressed in SM.

Since the only kinds of expressions of SM for which names appear in SM are sentences, it should come as no surprise that the only syntactical properties that can be expressed in SM are properties of sentences. For example, we can not say in SM that a certain expression is a matrix with three free variables. This does not seem like such an undesirable limitation, however, when it is kept in mind that modal predicates are supposed to express properties of sentences. That is, SM is intended to be a formal theory about sentences and it in fact turns out to be about sentences.

It does not seem unreasonable to expect that a syntactical theory of modality should be able to express the identity and non-identity of sentences. This is especially true for a theory which we claim would be sufficient to represent a modal antinomy if there was one.

If the first sentence on the blackboard B was identical with, or different from, the second sentence on B, we should be able to use this fact, if we wished, in trying to construct an antinomy.

Such identity relations can be expressed in SM. Whenever P and Q are the same sentence of SM we can prove $\ulcorner$'P' = 'Q'$\urcorner$, and whenever P and Q are different sentences we can prove $\ulcorner$'P' $\neq$ 'Q'$\urcorner$.

To be able to express truth functional relations between sentences also seems desirable for a syntactical theory of modality. One of the purposes of a theory of modality is to show how different modal predicates are related to one another. The various modal predicates are all interdefinable by the use of truth functional relations; impossibility being the necessity of the negation, entailment being the necessity of the conditional, and so on.

Whether or not a sentence is related to another sentence or sentences in a given truth functional way is decidable in SM. If P is the conditional whose antecedent is Q and whose consequent is R, we can prove $\ulcorner$'P' = cond('Q','R')$\urcorner$ in SM; and if P is not that conditional we can prove $\ulcorner$'P' $\neq$ cond('Q','R')$\urcorner$.

Thus, it is far from clear that SM is syntactically deficient, if this means failing to meet some requirement

which it is obvious that any syntactical theory of modality should meet. SM is quite a powerful theory. This is significant when it is remembered that Montague's Anti-Syntacticalist Thesis, quoted in Section Four, held that it was impossible to have an adequate syntactical treatment of even so weak a system as Lewis' S1.

/14/

SOME PROBLEMS INVOLVED IN EVALUATING THE WEAK FORM OF THE PARITY OF NAMING THESIS

In this section we shall discuss some problems involved in evaluating the weak form of the Parity of Naming Thesis. Here we are taking the weak form of the Parity of Naming Thesis to be the claim that any formal theory which meets the antecedent conditions of the strong form of the Parity of Naming Thesis (6.3), together with the added condition that the theory contain a "reasonable" amount of syntax, is inconsistent.

Since SM meets the antecedent conditions of the strong form of the Parity of Naming Thesis, the weak form of the thesis implies that there is an, as yet unspecified, amount of syntax such that if SM is extended to contain that amount, then the resulting extension will be inconsistent.

If we asked how SM might be extended so as to contain more syntax, the obvious suggestion would be that SM should be augmented by some "concatenation" theory. This would involve extending SM so that the resulting theory contained a name for each symbol of itself and a two-place concatenation function which would name expressions by concatenating the names of the symbols of which the

expressions were concatenates.

Suppose that "*" is a two-place function letter of L which we agree to interpret as the concatenation function. We also agree to abbreviate ⌜*(t,u)⌝ by ⌜t*u⌝ and to ignore the parentheses due to iteration.

If we adopt the method of TARSKI (1931) we could introduce names for symbols of L as follows. "En" names "N", "left" names "(", "aye" names "a", and "right" names ")", assuming that the new names we have chosen are all among the constants of L.

Consider the sentence "N(a)" of SM. The name of this sentence would be "En*left*aye*right". It would be said to be necessary by the sentence "N(En*left*aye*right)".

Tarski's method, although used in the typical concatenation theory, does not suit our purposes. We do not want to limit the number of times that the predicate "N" may be iterated. To avoid a limit on Tarski's method we would need an infinite hierarchy of names. If we wanted a name for the sentence "N(En*left*aye*right)" we would need names for "En", "*", "left", "aye", and "right".

To avoid the awkwardness of an infinite hierarchy of names the Syntacticalist should employ a naming function. The single quotation marks of L can be used for this purpose if a slight alteration is made in the specification

of what counts as a term of L. The single quotation marks must now be used to form names only of symbols of L. That is, "'N'" will be the name of "N" and "''''" will be the name of the single quotation mark. The name of the sentence "N(a)" will be "'N'*'('*'a'*')'".

By judicious use of abbreviations, SM with concatenation theory may be made to look like the unexpanded SM. We may use $\ulcorner$'st'$\urcorner$ as an abbreviation for $\ulcorner$'s'*'t'$\urcorner$. That is, "'N(a)'" may abbreviate "''''*'N'*'('*'a'*')'*''''". Also, we may use $\ulcorner$cond(t,u)$\urcorner$ to abbreviate $\ulcorner$t*'$\longrightarrow$'*u$\urcorner$.

We must now ask how much syntax is needed to produce an inconsistent extension of SM. To answer this question we should take advantage of the proof of Montague's Theorem in Section Five by trying to construct a similar proof for an extension of SM.

We define the diagonalization of a sentence with respect to the individual constant "a" to be the result of replacing all free occurrences of "a" in the sentence by the name of the sentence. The "a-diagonalization" of "N(a)" would be "N('N(a)')".

A fundamental principle of the Syntacticalist's position is that necessity can be meaningfully predicated only of sentences. Using a-diagonalization rather than the customary diagonalizations with respect to free variables guarantees that diagonalizations are sentences and

may meaningfully be said to be necessary.

Suppose that A is a finite conjunction of axioms for concatenation theory and that A is among the non-logical axioms of an extension of SM. Suppose further that there is a two-place predicate D(x,y) in this extension such that $\ulcorner(x)(\ D('P',x) \longleftrightarrow x = 'Q'\ )\urcorner$ can be derived from A whenever Q is the a-diagonalization of P. The extension is inconsistent if these conditions are met.

To show that this extension of SM is inconsistent by a proof parallel to that of Section Five we should have to find a sentence S, which is derivable from A, and which is such that

(14.1) $\vdash\ S \longleftrightarrow N('A \longrightarrow -S')$.

Consider the sentences

(14.2) $(Ex)(Ey)(\ N(x)\ \&\ x = cond('A',neg(y))\ \&\ D(a,y)\ )$,

abbreviated as "K", and

(14.3) $(Ex)(Ey)(\ N(x)\ \&\ x = cond('A',neg(y))\ \&\ D('K',y)\ )$.

It is easy to show that (14.3) qualifies as the sentence S of (14.1). The rest of the proof then proceeds as in Section Five.

There are two questions that must now be asked. The first is whether there is such an amount A of syntax. The second is whether A, if it exists, is a reasonable amount of syntax.

The answer to the first question is in all probability affirmative. Although we did not actually develop a system of concatenation theory in any detail we know that, in general, concatenation theories are quite powerful. QUINE (1946), for example, shows that arithmetic can be constructed on the basis of concatenation theory, implying that concatenation theory is at least as powerful as arithmetic. We know that finite axiomatizations of arithmetic are sufficient to represent all recursive functions, so it is to be expected that a finite axiomatization of syntax is sufficient to represent the a-diagonalization function.

We turn to the question of whether or not A is a reasonable amount of syntax. Suppose that the Syntacticalist claims that A is an unreasonable amount of syntax while the Anti-Syntacticalist insists that A is quite reasonable. How is this dispute to be settled?

We first look at the matter from the position of Montague and the Anti-Syntacticalists. At this point it looks as though they have the superior position. What more is needed to show that A is reasonable than to have

it admitted that A is true? Surely if the conjunction of A and the self referential instances of certain modal schemata are inconsistent, it is the latter that should be abandoned. Are there any other cases of sensible theories which become inconsistent when extended by the addition of true statements? It seems clear that the burden of proof is upon the Syntacticalist to provide some general principle upon which A might be judged to be unreasonable.

There are numerous presentations of concatenation theories in the literature. QUINE (1951), MONTAGUE (1957), MONTAGUE (1959), and MARTIN (1958) present axiomatizations for such theories.

There does not seem to be any point to our giving a detailed presentation of some one of these versions of concatenation theory. For if the Syntacticalist is to make good his case, he must argue that all such theories are somehow unreasonable in virtue of certain general features that they all share.

One common feature of the theories mentioned above, on which the Syntacticalist might well seize, is that in all of them the concatenation function is extensional, or referentially transparent. That is, statements such as

(14.4) $(w)(x)(y)(z)(w = x \longrightarrow y*w*z = y*x*z)$

are theorems of those theories.

For the languages whose syntax the authors are discussing (14.4) seems to be desirable as a theorem. But languages whose syntax they discuss are straightforward first order languages, which L is not. L contains the single quotation mark as a non-standard knid of function which forms terms, not from other terms, but from expressions which are not terms.

Our intention in constructing the language L was to construct a formal language in which the behavior of quotation marks in ordinary English was simulated. A characteristic feature of quotation marks in English is that they are intensional, or referentially opaque. If (14.4) was a theorem of an extension of SM, as it would be if concatenation were treated in the usual way, we would have the following as an instantiation of (14.4).

(14.5) $a = b \longrightarrow \text{'''}*a*\text{'''} = \text{'''}*b*\text{'''}$

On the intended interpretation of SM the individual constants are to be interpreted as expressions which refer to sentences. (14.5) when so interpreted might be read as

(14.6) If the first sentence on B is identical with the second sentence on B, then "the first sentence on B" is identical with "the second sen-

tence on B",

which seems to be an obvious falsehood.

The Syntacticalist could claim that the amount A of syntax needed to produce a contradiction in an extension of SM would be unreasonable if it led to theorems like (14.5).

It seems that the burden of proof has been shifted. No current version of concatenation theory prohibits statements like (14.5) from being theorems. It is incumbent upon the Anti-Syntacticalist to show that such a theory of syntax can be devised.

The Syntacticalist can go on to argue that the presence of theorems like (14.5) is not just an accidental shortcoming of concatenation theories, but rather an essential defect which illustrates his initial objection to the vague form of the Parity of Naming Thesis.

The vague form of the thesis, as stated in Section Six, was the claim that results equivalent to those for formal theories which named expressions by a system of Godel numbering could also be obtained for theories which named expressions by means analogous to the use of quotation marks. The Syntacticalist's objection was that the use of quotation marks might well be quite different.

The Syntacticalist can point out that it is just in

terms of their referential opacity that quotation marks differ from the numerals used in a Godel numbering. Quotation marks may have natural safeguards against the derivation of a contradiction which the numerals lack.

Although we have not given a definite answer to the question of the truth or falsity of the weak form of the Parity of Naming Thesis, we can make some remarks upon the relation between the weak form of the thesis and the question of the existence of a modal antinomy.

Suppose that the weak form of the thesis is true. It does not seem that this in itself would be a clear verdict in favor of the Antinomist. He would still have to show that the syntactical truths embodied in the reasonable amount of syntax needed to show that certain formal theories are inconsistent are also syntactical truths that can be used in the derivation of a contradiction from instances of modal schemata and statements about the arrangement of sentences on some blackboard.

There is no a priori reason to think that the Antinomist can always do this. The syntax of a formal language may be quite different from the syntax of ordinary English. For example, a not inconceivable state of affairs is that the syntax of a formal language might be finitely axiomatizable, this fact being essential to the proof that certain theories in that language are inconsistent,

while the corresponding amount of syntax for ordinary English might not be finitely axiomatizable.

The situation seems different, however, if the weak form of the thesis is false. This would count quite heavily in favor of the Non-Antinomist.

The Antinomist suffered a setback with the proof that every I-extension of SM was consistent. The Antinomist then turned towards the weak form of the Parity of Naming Thesis as a heuristic guide. He had no clear idea of what the extra assumptions needed to derive a contradiction might be, but he hoped that a proof of the weak form of the thesis would provide some insight. If the weak form of the thesis is false there does not seem to be much hope for the Antinomist's position.

BIBLIOGRAPHY

Page references to works mentioned in this bibliography are to the first source cited for that work. The ordering of sources is often determined by accessibility rather than chronology. The parenthesized dates are intended to correspond more nearly to the original presentation of the material than to the publication date of the first mentioned source.

ANDERSON, Alan Ross

(1965) "Does Negation Rest upon a Mistake?" Philosophical Studies, v. 16, 1965, pp. 4-6.

ANDERSON, Alan Ross and BELKNAP, Nuel D.

(1959) "A Simple Proof of Godel's Completeness Theorem" (Abstract) The Journal of Symbolic Logic, v. 24, 1959, p. 320.

FEYS, Robert

(1965) Modal Logics Edited with some complements by Joseph Dopp, E. Nauwelaerts, Louvain, 1965.

GODEL, Kurt

(1933) "Eine Interpretation des Intuitionistischen Aussagenkalkuls" Ergebnisse eines Mathematischen Kolloquiums, v. 4, 1933, pp. 39-40.

KAPLAN, David and MONTAGUE, Richard

(1960) "A Paradox Regained" Notre Dame Journal of Formal Logic, v. 1, 1960, pp. 79-90.

KNEALE, William and KNEALE, Martha

(1962) The Development of Logic The Clarendon Press, Oxford, 1962.

KRIPKE, Saul A.

(1959) "A Completeness Theorem in Modal Logic" The Journal of Symbolic Logic, v. 24, 1959, pp. 1-14.

LANGFORD, Cooper H. and LEWIS, Clarence I.

(1932) Symbolic Logic The Century Company, New York, 1932.

MARTIN, Richard M.

(1958) Truth and Denotation Routledge & Kegan Paul Ltd., London, 1958.

MENDELSON, Elliot

(1964) Introduction to Mathematical Logic D. Van Nostrand Company, Inc., Princeton, 1964.

MONTAGUE, Richard

(1957) Contributions to the Axiomatic Foundations of Set Theory Unpublished doctoral dissertation, deposited in the University of California library, Berkeley, 1957.

(1959) "Semantical Closure and Non-Finite Axiomatizability" Infinitistic Methods, Pergamon Press, New York, 1961, pp. 45-69.

(1963) "Syntactical Treatments of Modality with Corrollaries on Reflexion Principles and Finite Axiomatizability" Acta Philosophica Fennica, v. 16, 1963, pp. 153-67.

MOSTOWSKI, A., ROBINSON, R.M. and TARSKI, Alfred

(1953) Undecidable Theories North-Holland Publishing Co., Amsterdam 1953.

PRIOR, A.N.

(1963) "Is the Concept of Referential Opacity Really Necessary?" Acta Philosophica Fennica, v. 16, 1963, pp. 189-99.

QUINE, W.V.

(1946) "Concatenation as a Basis for Arithmetic" The Journal of Symbolic Logic, v. 11, 1946, pp. 105-14.

(1951) Mathematical Logic Revised edition, Harvard University Press, Cambridge, 1951.

(1953) "Three Grades of Modal Involvement" Proceedings of the XIth International Congress of Philosophy, v. 14, North-Holland Publishing Company, Amsterdam, 1953, pp. 65-81. The Ways of Paradox, Random House, New York, 1966, pp. 156-74.

(1961) From a Logical Point of View Second edition, revised, Harvard University Press, Cambridge, 1961.

SMULLYAN, Raymond M.

(1968) First Order Logic Springer-Verlag New York Inc., New York, 1968.

STRAWSON, P.F.

(1948) "Necessary Propositions and Entailment Statements" Mind, v. 57, 1948, pp. 184-200.

(1949) "truth" Analysis, v. 9, 1949, pp. 83-97.

TARSKI, Alfred

(1931) "The Concept of Truth in Formalized Languages" Logic, Semantics, Metamathematics, translated by J.H. Woodger, The Clarendon Press, Oxford, 1965, pp. 152-278

(1944) "The Semantic Conception of Truth" Readings in Philosophical Analysis, edited by H. Feigl and W. Sellars, Appleton-Century-Crofts, Inc., New York, 1949, pp. 52-84. Philosophy and Phenomenological Research, v. 4, 1944, pp. 341-75.

ZIFF, Paul

(1960) Semantic Analysis. Cornell University Press, Ithaca, 1960.