HILBERT, DETLEFSEN AND A FREGEAN RESPONSE

by

RUDY HENRY VOGT

B.A., The University of British Columbia, 1986

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in
THE FACULTY OF GRADUATE STUDIES

Department of Philosophy

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1993

© Rudy Henry Vogt, 1993
In presenting this thesis in partial fulfilment of the requirements for an advanced
degree at the University of British Columbia, I agree that the Library shall make it
freely available for reference and study. I further agree that permission for extensive
copying of this thesis for scholarly purposes may be granted by the head of my
department or by his or her representatives. It is understood that copying or
publication of this thesis for financial gain shall not be allowed without my written
permission.

(Signature)

Department of Philosophy
The University of British Columbia
Vancouver, Canada

Date 4/30/1993
ABSTRACT

HILBERT, DETLEFSEN AND A FREGEAN RESPONSE

This thesis examines Frege’s objections to Hilbert’s program. It argues that Frege’s concerns can best be understood as questioning Hilbert’s implicit importing of content into ideal mathematics. A contemporary defense of Hilbert, Michael Detlefsen’s *Hilbert’s Program*, is taken as representative of an anti-Fregean view. Detlefsen, following Hilbert, develops the finitary/ideal distinction in mathematics. Finitary mathematics, unlike ideal mathematics, is claimed to involve genuine propositions. Frege’s objections are thus primarily directed against ideal mathematics. More accurately, Frege’s objections are seen as demonstrating the need for explicitly identifying the background language in which Hilbert’s program, and thus model theory, are carried out.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER ONE</td>
<td>5</td>
</tr>
<tr>
<td>CHAPTER TWO</td>
<td>37</td>
</tr>
<tr>
<td>CHAPTER THREE</td>
<td>66</td>
</tr>
<tr>
<td>CHAPTER FOUR</td>
<td>84</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>101</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>103</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I am grateful for the opportunity to express my indebtedness to the following people. I would like to thank Andrew Irvine, who not only encouraged me and made valuable suggestions, but without whom this thesis would not have been completed, and Jack Stewart for his patience, and willingness to listen and comment on my arguments. I am grateful to Dick Robinson, who took time from his marking to read this thesis and comment on it.

Mark Glouberman and John Klippenstein over the years provided me with invaluable suggestions and criticisms. Their curiosity and joy of pursuing philosophical debate was both enriching and enjoyable. Joan Bryans read the third chapter and provided valuable suggestions and encouragement. Joan Bryans, Peter Apostoli, John McGuire and Tim Kosub, on different occasions provided me with hours of enjoyable debate and friendship. Peter McKinnon, not only patiently listened to and engaged my arguments but also solved all my computer problems.

To my parents Rudolf and Margerete Vogt, and all my family, who not only provided support but also showed confidence in me I am very thankful. I would also like to thank Alexander and Margaret Ferguson for the support they provided. I would like to thank Helmut Buhr, who over the years listened to my arguments and who on more than one occasion provided solace and encouragement.

Finally I would like to thank Janet Vogt for her unwavering confidence in me and the encouragement that she showed me in my pursuit of this goal. For Janet and the kindness and caring she showed me, I am deeply grateful.
INTRODUCTION

Following Hilbert, Michael Detlefsen argues that one characteristic trait of finitary reasoning is its reliability. Another is that it is meaningful, or that it involves genuine propositional content. Finitary mathematics therefore operates upon real or contentual propositions, the content of which provides epistemic value. In contrast to finitary mathematics, ideal mathematics does not involve genuine propositional content. Ideal mathematical statements are best understood as syntactic elements whose content does not exceed the order and regularity of the symbols themselves. Unlike ideal proofs, finitary proofs are characterizable by how the content of a finitary statement is effected by preceding contentual statements. Finitary mathematical reasoning is therefore often referred to as real or contentual reasoning, in virtue of the propositional content that is being operated upon. Strictly speaking, statements of ideal mathematics are meaningless. In spite of this fact, ideal proofs can be evaluated as to their reliability. Their reliability is a function of the terminal formula and whether or not it expresses a finitary truth.

An important question arises in attempting to understand how this evaluation works. Detlefsen argues that the evaluation relies only upon finitary reasoning. Clearly, "finitary reasoning" cannot then be explained by appealing to the content of those formulas being evaluated, since ideal mathematics is

\[1\] Detlefsen, 1986, pp. 6f.
contentless. Instead, the evaluation of an ideal proof consists in assessing the terminal formula to be a finitary truth, and in assessing the form of an ideal proof as being of a reliable type. Detlefsen argues that finitary reasoning is not constrained by "inference free" computations. Since finitary reasoning makes use of schematic letters and proof-schemas, both of which involve inferences, finitary reasoning is essentially reasoning on types.

Frege's concern is that ideal mathematics must appeal to a kind of reasoning which is disjoint from finitary reasoning.\(^2\) Thus, according to Frege the purported relationship between ideal and real mathematics is left unexplained. In order to counter Frege's objections, Detlefsen attempts to show that the evaluation of ideal proofs involves nothing more than finitary reasoning applied to types of proofs, and thus reliability is preserved even in ideal proofs.

In this thesis I intend to show that in Detlefsen's characterization of the relationship between finitary and ideal mathematics, the content of finitary mathematics plays an increasingly diminished role. Detlefsen is forced, either to accept this view or assign content to ideal proofs. Frege's objections focus on the Hilbertian's inability to give an account of his own project without appealing to the taboo content denied him.

\(^2\) While Frege does not explicitly state this view, we can deduce it from the following two claims. The first is that all arithmetical inferences can be reduced to logical inferences. (As I will argue in Chapter Two, for Frege logical inferences involve conceptual analyses.) The second claim is that Frege does not hold the manipulation of meaningless symbols as a species of logical inference. See Frege, 1885, pp. 142f. & 145.
In what follows, Chapter One performs two important functions. The first is to provide an explanation of Detlefsen's account of Hilbert's program. The second is to argue that Detlefsen's account of Hilbert's division of mathematics into finitary and ideal can only avoid Frege's criticism - that ideal mathematical reasoning is disjoint from contentual reasoning and is thus unmotivated - by diminishing the role that the contentual plays in finitary arithmetic. Furthermore, it is argued that the diminishing role of content threatens to relativize the concept of "being finite", and thus weakens our initial confidence in even finitary reasoning.

In Chapter Two Frege's objections to Hilbert's program are introduced. Frege's critical remarks concerning Hilbert are generally understood as representing a misunderstanding of the role the uninterpreted symbols play for Hilbert. It will be argued that one need not understand Frege's objections as a misunderstanding of the function of the uninterpreted symbols. Rather, we should understand it as an unwillingness to take the formalist account as correctly representing what is actually taking place. This unwillingness to see the formalist account as a reliable account of the "doing of mathematics" suggests an alternate role for Frege's formalism than is sometimes given. Frege's formal language is thus understood to be an idealization of the content of a natural language. As such, the question of content is always predominant in Frege's critical remarks to Hilbert. Frege's two main arguments for why Hilbert's independence proofs are irrelevant are examined. It is argued that Frege's characterization of logic as a species of conceptual analysis reflects a worry that the Hilbertian formalist ignores. The tension between
finitary and ideal statements is to be seen as asking for justification of the ideal statements as extensional elements. That some extension of finitary mathematics provides a reliable solution to the problem does not entail that the extension is appropriate.

In the Chapter Three, Frege’s concerns are reintroduced. Frege’s objection, that content must implicitly be imported in order to understand how uninterpreted symbols function, is examined with respect to model theory. Specifically, Hodges’, 1985/86, account of how Frege misunderstood the model theoretic move in Hilbert’s work is evaluated.

In Chapter Four Detlefsen’s conflation of a metamathematical formula and its propositional content will be examined. It will be argued that Detlefsen’s characterization of Gödel’s incompleteness theorem, which involves characterizing the conditions for the unprovability of an object language formula in a theory, T, cannot be generalized to provability in T. I argue that at best we can show that a meta level formula has been proved. Detlefsen, in attempting to show how the consequences of Gödel’s incompleteness theorem can be avoided, raises doubts about the relationship between the syntactic formula and its propositional content. Following a Fregean argument this division is examined and applied to the meta level formulas and their content. It is this move which engenders doubts about whether Detlefsen can coherently give an account of what it is for a metamathematical proposition to be proved.
CHAPTER ONE

In *Hilbert’s Program*, 1986, Michael Detlefsen defends a Hilbertian thesis against Frege’s attacks. Detlefsen’s defence of Hilbert’s thesis is both an attempt to articulate where Hilbert’s position is defendable and to bolster Hilbert’s position where it is not.

Detlefsen begins his defence of Hilbert by developing a version of instrumentalism to which he takes Hilbert to be committed. Instrumentalism, as Detlefsen describes it, is applicable only to part of mathematics, namely that part of mathematics which he calls ideal mathematics. Instrumentalism with respect to some body, T, of theorems and proofs consists in

the belief that the epistemic potency of T (i.e., the usefulness of items of T as devices for obtaining valuable epistemic attitudes toward genuine propositions of some sort) can be accounted for without treating the elements of T literally (i.e., as genuine propositions and proofs), but rather as "inference tickets" of some sort. (Detlefsen, 1986, p. 3)

It is with respect to ideal mathematics, that is, that part of mathematics which takes advantage of non-finitary formulas, that Hilbert is an instrumentalist. Clearly Detlefsen must give an account of the genuine, finitary propositions and proofs of mathematics in contrast to those of ideal mathematics. Detlefsen diffuses this obligation to some extent by locating the ontological commitments of mathematics only in that part of mathematics which establishes the *reliability* of the mathematics being used. (Detlefsen, 1986, p. 3) This commitment creates a division between the genuine propositions and the non-genuine propositions. From
the instrumentalist perspective, the judgement that some tool is reliable is a judgement which looks towards the final result, and to the ability of the instrument to duplicate this result reliably whenever and wherever the instrument is used. Thus instrumentalism is characterized by pragmatic conditions. Whether or not our instrument is reliable is to be understood as a function of the indirect truth conditions of instrumentalism. These indirect truth conditions are not to be understood as being independent of our assessment of a tool as being reliable. Yet the degree to which these conditions have been tailored to accommodate non-genuine propositions threatens to undermine the very distinctiveness of the genuine propositions themselves. Propositions are characterized by their content. The need to find some common characterization for both genuine and non-genuine propositions threatens to weaken the ontological commitment of the genuine propositions. The judgement that some tool is reliable is a judgement that it is reliable in some kind of situation. The instrumentalist can discharge his obligation to provide an account for the content of the non-genuine propositions only to the degree that he is able to characterize the "kind of situation" in which a particular tool is reliable. Detlefsen must be able to characterize the "kind of situation" in which a tool is reliable without appealing to some content that non-genuine propositions implicitly express. Furthermore, if the content needed to characterize this "kind of situation" is unique to the genuine propositions, then the characterization of the unity of ideal and real mathematics must explain the uniqueness of genuine propositional content. It might appear as if this obligation
is easily discharged. After all, the characterization of the project for which a tool is reliable is in a large part dependent upon the tool itself. As an example, it might be characterizable outside of the project with reference to the general domain of mathematics. Yet, however appropriate these extra-mathematical conditions might be for tools that are not mathematical, this explanation is not appropriate for mathematical tools. It is not appropriate for at least two reasons. One is due specifically to Hilbert’s own concerns; the other is more general.

Consider first Hilbert’s reason. Hallett argues that Hilbert’s holism incorporates all possible facts of importance; any fact which is relevant is to be included in the theory. (Hallett, 1990, p. 200) Holism raises a number of important questions. In the first place it is unclear how the object language and metalanguage are to be related. If our metalanguage is understood to be an informal language (Detlefsen explicitly refers to it as being informal on a number of occasions) and the relevant relations which implicitly provide the definition of object language terms are those of the formalized object language theory, then the metalanguage appears to be unhelpful for understanding terms of the object language theory itself. Reference to propositions in the metalanguage looks illegitimate, since the object language provides the appropriate conditions for understanding what the terms mean. If, on the other hand, the relevant relations are those in both the metalanguage and the object language theory, then the danger of inconsistency in the metalanguage is great enough to provide concern. The fact that there is an informal aspect to the metalanguage implies that, if there
is a contradiction present, it may not be discovered immediately. This problem is accented in the case of Detlefsen’s defence of Hilbert’s program. The relationship between the metalanguage and the object language is not 1-1. Our metalanguage propositions may turn out to be incoherent. If our confidence in the consistency of the metalanguage is not very high, then since these metalanguage propositions can be expressed in different ways in the object language, and since only some of these expressions are provable in the object language, we must accordingly be convinced that the metalanguage is consistent in order for us to be confident that the propositional content of a provable object language formula is relevant. If the metalanguage is inconsistent, then the proof of an object language formula, expressing a metalanguage proposition, can weaken our confidence in our theory.

Holistic definitions depend upon all the relations that a term has in a theory. Thus, in order to know what the term "point" means, one must know the relations this term has to all other terms in that theory. Hallett points out that completeness was an important condition on theories for Hilbert because it "fixed" the meanings of the terms in a theory. (Hallett, 1990, p. 201) Recursive definitions, in a first-order language, are constructive. The attempt is to construct some set of objects such that this set is characterized by a specific recursive property. In a second-order language we can treat this definition non-constructively. Yet irrespective of whether the recursive definitions are constructive, these definitions do not depend upon the relations of all the terms in the theory.
A related problem for the holist is that there is no "outside" from which extra-mathematical conditions can be imported to provide the appropriate conditions which might characterize the "kind of situation" in which the tool is supposed to be reliable. That a mathematical tool is reliable is supposed to be self-evident with respect to its formal properties and not dependent upon some independent characterization of its use. If our understanding of "reliability" is dependent upon the resultant formula conforming to an already understood formula, then our understanding of "reliability" could always be seen to be dependent upon the kinds of "things" which the resultant formulas characterize. As such, the concept of "reliability" would be rendered defeasible by other commitments to our mathematical ontology.

Detlefsen's objection is that "reliability" is to be understood in terms of the finitary procedures characteristic of finitary arithmetic. As such, ideal mathematics is to be conservative over finitary results, not over the objects on which finitary reasoning operates. The task is to explain our understanding of "reliability" without appealing to the content on which finitary mathematical procedures operate. The degree to which these arithmetical procedures guarantee "reliability" due to their content is the degree to which ideal mathematics fails to be analogous to finitary arithmetic. There is a strong motivation in Detlefsen's writing to understand the content of real mathematics as consisting in the form or shape of its tokens. In much of what follows in this chapter I argue that this explanation should be resisted. Detlefsen cannot whole-heartily adopt this view without adopting the
view that either we attribute instrumentalism to real mathematics or that ideal mathematics incorporates meaningful expressions. Clearly, if the resultant formula of a proof is suitably finitary and short, then the reliability of the computation will be self-evident. That is, it will be self-evident that the resultant formula is in fact the correct kind of formula, without any appeal to a prior understanding of what a correct formula here means.

Kreisel argues that a computational proof can be said to avoid any use of inference. Since computations are subject to actual concrete manipulations, any inference in a computational proof reduces to a transformation of one symbol to another. Our confidence in the reliability of a computational proof is thus related to the proof's self-evidentness. That no special faculty is required obviates the need for inferences. Detlefsen, in contradistinction, argues that even computational proofs require minimal inferences. The minimal inferences involved are just the inferences necessary to represent the numerals as having a similar form or structure such that the schema of a proof is potent enough to perform its expressed task. It is this move that allows Detlefsen to argue that the proofs of ideal mathematics are analogous to the proofs of contentual mathematics.

As Philip Kitcher points out, the attack on Hilbert's program should be directed against the very reliability of the tools of the mathematician, since these very tools have led to uses which were found to justify claims which are in fact false. (Kitcher, 1976, pp. 99f.) Detlefsen's evaluation of ideal mathematics using

---

1 As characterized by Detlefsen, 1986, p. 52.
contentual mathematical judgments is an attempt to constrain the domain of procedural reliability to that of the contentual domain without diminishing the importance of ideal computations. The degree to which the reliability of contentual mathematical formulas must be seen to be reliable due to their content, and the degree to which content influences the procedure, would then gauge the degree to which ideal mathematics is disanalogous to finitary mathematics.

The second more general problem for Hilbert's holism is just that the characterization of the reliability of some tool, e.g. a hammer as being reliable for building a house out of wood and nails, is not diminished in cases where one is not able to characterize all the kinds of situations in which that tool might be used. The practical purposes for which a hammer might be used outstrip the typical characterization of a hammer's reliability. Mathematical tools, on the other hand, in principle involve a kind of completeness of characterization which, while perhaps not unique to mathematics, is extremely important to it. What is important is that a mathematical tool is reliable in any procedure in which it can be used and, if it cannot be used, then it is clear, in that procedure, that it is not reliable. The point is that the tools in question necessitate a kind of closure on their use with respect to reliability. Thus any mathematical tool whose reliability is not questioned will provide a result of the right kind, namely a mathematical result of equal reliability in any of its uses. But in order to be confident of this reliability, we must be confident that the domain, the ontology of mathematics, can in fact be characterized adequately without begging the question. Only then will it be evident
that the failure of a tool’s reliability is not due to a mistake. The kinds of mistakes relevant are those of human error in computation and those of using a procedure which is not supposed to be adequate. Without the condition of closure on the procedure of the proofs, the expressed strength of even finitary proofs will show a net loss of confidence in them. Computational proofs, as characterized by Kreisel, can be seen as avoiding this problem by the avoidance of actual inferences, since inferences demand evaluations.

Detlefsen characterizes instrumentalism as involving some set of symbols which are to be evaluated, not in terms of their semantic content, but as inference tickets. Inference tickets function by endowing the terminal formula with truthfulness, without committing us to any specific epistemic attitude regarding the truthfulness of the inference ticket itself. Genuine proofs exploit judgments regarding the premises and inferences as true or truth-preserving, while ideal computations provide

for the epistemic acquisition of a proposition $P$ to which it leads by exploiting a judgement regarding its (the computation’s) formal character and a meta-computational judgement affirming the reliability of computations having that formal character as guides to truths of the type exemplified by $P$. (Detlefsen 1986, p. 4)

Proofs in general can therefore be characterized by the following three conditions:

\footnote{This characterization is incomplete for it does not yet consider Detlefsen’s condition that a proof must conform to an accepted kind of finite character. That is, the evaluation of a proof at the metalevel is performed by using finitary criteria.}
(1) Propositions used in mathematical proofs are divided into real (or meaningful) propositions, and ideal (or non-genuine) propositions.

(2) The pre-terminal formulas and inferences of an ideal proof are not meaningful or truth-preserving. Instead, their epistemic utility arises as a result of their metatheoretic formal properties.

(3) The metalevel judgments must be meaningful.

Evidence that Hilbert is in agreement with these conditions is given by Detlefsen in the form of in a long quotation taken from "On The Infinite". (Hilbert, 1926, pp. 195f.)³ Before proceeding with Detlefsen's argument, it is worthwhile examining this quotation from Hilbert. In the quotation Hilbert makes an initial distinction between intuitive number theory and elementary number theory. It appears that intuitive number theory deals with the elements of arithmetic directly. Formulas are used only for communicating and letters stand for specific numerals. Elementary mathematics, on the other hand, takes the formulas to be independent objects. The contentual propositions of number theory are then formalized by means of these independent objects. (Detlefsen, 1986, p. 5)⁴ Algebra, unlike number theory, does not deal with real or contentual sentences. The formal object is related to the contentual proposition by substitution. The formula

³ Detlefsen’s version of this quotation is from the translation given in van Heijenoort [1967b] and exhibits some differences from that given in Benacerraf & Putnam [1983].

⁴ The difference between the van Heijenoort [1967b] translation and the Benacerraf & Putnam [1983] translation, while not relevant here, is quite large.
a + b = b + a

is not meaningful. Only when we uniformly substitute, for example the numerals "5" and "7" for "a" and "b" respectively, is the new formula meaningful. Nevertheless Hilbert’s exposition appears to be incomplete. Later he states that the above formula is meaningless and he mentions that "a" and "b", as well as "=" and "+", are without content. The substitution rule must apply to "+" (standard meaning), and to "=" (standard meaning), as well as to some specific numerals, say "5" and "7", which give the meaningless formula "a + b = b + a" content.

It might seem that my contention that we take the substitution rule in its wide scope (in which everything is substitutable) is irrelevant. As long as the terminal formula is always true, the computation will guarantee reliable results only in virtue of its formal character. Frege argued that without some parameters (e.g. keeping the meaning of the logical constants fixed) our translations become subject to an unexamined conceptual framework. This framework takes concepts to exist independently of the logician’s decision to regard our understanding of a formula as a function of specific interests.

John Etchemendy has argued that when these parameters are violated our substitution rules provide counter-intuitive results. (Etchemendy, 1988, pp. 98ff.) Thus if we define a "logical truth" as any truth which is unaffected by content, and if we allow all the symbols in the sentence

\[ \text{5 I will refer to the "substitution rule", when it is applied not only to "a" and "b" but also to "+" and "="}, \text{as being applied in its wide scope.} \]
(x)(fx v -fx)

to be open to substitution, then one would expect that we would discover the set of truths that is unaffected by the content of the symbols. Yet as Etchemendy points out, if we allow our logical constants to be open to substitution, then no sentence of our logical language will be a logical truth. Similarly, if we keep all the meanings of the symbols fixed (not allowing the substitution rule to apply to any symbol), then by default every sentence will result in a logical truth. Content is supposed to constrain the number of models in which a formula is true. Formulas that are true in spite of their content are logical truths, and logical truths are true in all models.

To the question of why one tool is better than another, one could argue that for any two tools, if the two resultant formulas are equivalent there is no preferred choice between them. If a black box produced the same terminal formula which our computations did, there would accordingly be no reason to choose between the black box and its internal configurations and our computations. Nevertheless, the mathematician is interested in mathematical tools. These interests are to be understood as placing requirements on the tools being used. One of these requirements is that if the scope of mathematics is extended, the entities that are introduced should in some relevant way conform to those already understood. Thus the introduction of imaginary numbers was seen as being important, because it allowed us to attain a system which was closed under the square root operation. That is, it performed a function by extending a concept over a larger domain. It did
this by keeping fixed certain meanings while either freeing others to be re-interpreted or allowing new entities to be introduced. Clearly these requirements are metalevel requirements, but this does not mean that they do not demand a certain toll from their intermediary steps, namely that the meaning of the square root sign is preserved in the intermediary steps.

The instrumentalist, as Detlefsen has begun to describe him, takes the terminal formula as contributing epistemic value in part in light of the formal properties of the calculations. Thus we would find the black box wanting if it did not mimic the requirements that Detlefsen places on the "formal". Whatever the properties that comprise the "formal", they must be constitutive of finitary arithmetic itself. In the case of the black box, if its internal configuration, is suitably finitary and arithmetical, its determination of the terminal formula would be acceptable. Hilbert, it seems, would disagree. Finitary mathematics deals with the numerals themselves. This is evident in the quote that Detlefsen mentions. Intuitive number theory differs from algebraic calculations in respect to the objects that are being manipulated. However one understands the move from concrete space-time tokens to the forms of contentual mathematics, the abstraction that takes place is clearly conservative over content. The abstraction preserves meaning. Without this preservation, the notion of "form" or "shape" would be syntactic. This is clearly not so in the case of ideal proofs. Since the symbols there are meaningless, there is no content to preserve.

Detlefsen characterizes Frege's concerns in the following manner.
Frege’s problem is that of explaining how the construction of an ideal proof \( I \) for a real (contentual) proposition \( R \) could conceivably be of any help in the fixing of an epistemically valuable attitude toward \( R \). Since \( I \) is not composed of genuine propositions, it cannot be seen as giving, in itself, a reason for \( R \) (which is what a genuine proof, in Frege’s sense, does). And if \( I \) doesn’t provide a reason for believing \( R \), how is it that the construction of \( I \) is relevant to the epistemic acquisition of \( R \) in a way that, say, doing the rumba or bowling isn’t? (Detlefsen, 1986, pp. 9f.)

The tension between Frege and Hilbert is then that

Hilbert’s instrumentalistic claim is that the ideal derivations are epistemically potent; Frege’s Problem is that he doesn’t see how they could be. (Detlefsen, 1986, p. 10)

Detlefsen is in agreement with Frege regarding the status of the premises of an ideal proof. Nonetheless, to judge an ideal proof on the premises’ inability to effect the terminal formula semantically is argued as being inappropriate. Ideal proofs are judged to be truth preserving by an evaluation at the metalevel. It is this evaluation that contributes the epistemic value to the judgement that the ideal proof is truth preserving. Ideal proofs are evaluated by determining

1. that the ideal proof is an object with a specific allowable syntactic character, and

2. that objects with this syntactic character have a terminal formula which expresses a true real proposition.

Determining, at the metalevel, that an ideal proof fulfills conditions (1) and (2) is insufficient to satisfy a Fregean. The Fregean wants to know how the overlay of an ideal proof with a syntactic template (or form) contributes to the judgement concerning the terminal formula. Returning to the black box, the requirement that
the internal configurations must fulfill is relatively benign. The black box need not be able to determine that its computations fulfill these requirements. As long as someone notices (evaluates) that the terminal outputs are consistently reliable, the further requirement that the internal configuration must have an appropriate syntactic character is trivial. It is irrelevant whether each step in the black box is motivated by the content of previous steps; each step is motivated by the formal characteristic of the previous step as having a specific kind of syntactic character. Intermediary steps need not preserve truthfulness. While the evaluation of the syntactic character of a proof at the metalevel is meaningful, the objects of the evaluation are not. The actual ideal language inscriptions are evaluated as being reliable only when some specific form or structure (template) can be applied to them. Implicitly, any mappings from the metalevel evaluation to the object level (ideal language) which preserves the intent of the metalevel evaluation will be sufficient. Since the subsequent steps in the black box are not motivated by any previous content, judgements are irrelevant. The metalevel evaluation that some specific syntactic element is of a specific kind relies upon the metalevel prover’s ability to assign a kind of order and template to the segment of symbols that are being examined. Thus whether steps 71 through 88 in the series of steps executed by the black box express the universal quantifier template cannot be denied or affirmed, since the evaluation is not conservative over content. The objects of evaluation and their content do not contribute to the question of which template to use. At best we can say that steps 71 through 88 and the universal quantifier
allow for some appropriate mapping. Ideal proofs are evaluated to be of a specific syntactic kind and proofs of this kind are reliable. Being a specific kind in the ideal language just means that there is a template which can be placed on some set of meaningless inscriptions. Instruments do not preserve content.

Typically Detlefsen settles on arguments which involve more than one computation. Yet it is very difficult to see how a syntactic characterization of a one step inference would function. Take for example the following two inferences:

(1) \( P \)

\( P \) or \( Q \), and

(2) \( \underline{P} \)

\( P \) and \( Q \)

Formally (1) and (2) can be seen as having the same structure. That is

(3) \( \underline{P} \)

\( P \) \& \( Q \)

is a possible template for both (1) and (2). None the less, (1) is a valid inference and (2) is not. One place inferences depend upon the connectives and the semantic content that they are meant to capture. One might object that (1) and (2) are finitary statements and not ideal. For any theory in which (1) or (2) are valid inferences they must be understood as either schemas or universal statements. The point is that the template with which we choose to overlay the two

\(^6\) A. N. Prior, 1960, makes a similar point in a related context.
subsequent formulas is dependent upon the semantic content that we are attempting to capture.

It is difficult to determine the relationship that a template or formal syntactic characterization has to intuitive arithmetic. The term "intuitive arithmetic" suggests that intuitions are of crucial importance in the inferences that are demanded from one step to another. Without these intuitions the demand for further explanation would be justified.

The black box's internal configuration is of finite length, but what makes this configuration dubious as a calculation is that we do not have an appropriate explanation of how such a calculation might be informed by our mathematical intuitions. Without these intuitions we are unable to know which of a myriad of templates one should overlay the internal configuration with. Something like a ceteris paribus clause is needed to determine which template is the appropriate one. If we further require that finite calculations be performed on numerals, the template, as applied to this configuration and as an independent criterion, becomes even less understandable. However one characterizes an inference rule on the numerals this rule seems to be informed by the nature of the numerals and the content they exhibit.

Tait understands the two concepts "finite sequence" and "iteration" to be primitive. (Tait, 1981, p. 532) Clearly these two operations are not the only two that are necessary. Repeatedly, in his explanations of a "finite sequence", Tait argues that we see (via a Kantian intuition) some specific sequences as finite, or
that for some finite iteration we can see that the recursive definition is applicable. Tait identifies this seeing as the primitive Kantian intuition necessary for doing arithmetic. It may appear that seeing a sequence as being finite is unproblematic for Tait. After all, although the inscription

(1) 1111111111111

can be seen to be an infinitely many different finite sequences, these different seeings all see (1) itself as being finite. But why should we concede that (1) is a finite sequence and that seeing (1) as finite involves no commitment to the "being finite" that Tait is attempting to explain. Take, for example, the following sequence:

(2) 11111111111...

At first glance (2) is seen to be different from (1), but this is so only because of the convention which interprets the "..." as representing an unending continuation. Implicitly, the last "111" of (1) is not equated with the "..." of (2). Clearly we do not see this, after all we could have established a convention which stated that in any iteration the last three "1"s will be seen as representing an unending continuation unless a "." was placed after them. The objection that placing a "." after an iteration involves a symbolic representation is surely beside the point. If we "see" (1) as some specific finite sequence, then we have already allowed some inscription to represent the value of being a "unit of iteration". Tait says of sequence (1) that we first see it as a finite sequence and that only then can we count and determine its number; numbers are specific determinations of finite
sequences. (Tait, 1981, p. 529) The determination of a sequence as a number is left completely unexplained. One reason why one might believe that "counting" and "finite sequences" are different operations is that the determination of an iteration of "1"s as finite is self-evident. The kind of intuition needed fulfills the conditions as given by Fraenkel et al. It is unclear that Tait can rely on this view. On the one hand, the fact that iterations of "1" are intuitively transparent does not mean that there is no intuition necessary for us to understand (2) as finite. On the other hand for some appropriately large n, in

(3) 111...n times...111

our determination that the iteration of "1" n + 6 times is finite relies upon our confidence that n is a finite number. The appeal to a decreasing order as in primitive recursive arithmetic does not explain how we can determine or see (3) to be a finite sequence. This follows since at some point this sequence is no longer surveyable and relies upon an assumption, viz. that the "1"s being iterated behave just like the "1"s examined (i.e. that there are only finite many of them). Tait must assume that for any finite iteration of "1"s our intuition of what occurs at the nth iteration is exactly like what happens at the second iteration.

Detlefsen's criticism of Tait's argument is that finitary proofs are not uniformly reliable. Due to the length and complexity of some finite proofs, the appeal to some "para-sensory" intuitions in order to explain the infallibility of finitary evidence fails. (Detlefsen, 1986, pp. 64f.) Detlefsen’s positive position is

---

7 These conditions will be examined shortly.
much less clear. Detlefsen seems to hold that the sorts of objects that we subject
to finitary constructions are of the "form" or "shape" of a sign-token.

The scope of a given finitary operation is, therefore, never just one
or another particular token, but rather the whole class of tokens
sharing a given form. (Detlefsen, 1986, p. 51)

For Detlefsen the very idea of "iteration" is problematic. Since the sign-token as
a concrete object is not the subject of our finite reasoning, there is, strictly
speaking, nothing to be iterated. Detlefsen's "intuition" takes the "whole class of
tokens" as already given. It is unclear how the concept of "surveyability" is to be
relevant to this class. Take for example the assumption that basic numerals may
be identified with stroke expressions. If, on the one hand, "surveyability"
demarcates the class of tokens that have a specific form or shape, then the "whole
class of tokens sharing a given form" cannot play the role in intuition that
Detlefsen seems to imply. If, on the other hand, "the whole class of tokens sharing
a given form" demarcates the class of tokens which have a specific form or shape,
then the surveyability of only some proper subclass is irrelevant. The class of
stroke expressions is already given. In either case the relevance of "iteration" and
primitive recursive arithmetic remains unmotivated in Detlefsen’s account.

Detlefsen argues that the evaluation of an ideal proof involves an
assessment of its form by finitary means. Finitary reasoning involves reasoning
over types. Evaluating ideal proofs therefore involves evaluating types of proofs.
Strictly speaking this characterization is false, since ideal symbols are meaningless.
Let us ignore Hilbert’s claim that finitary reasoning involves reasoning over objects
which are immediately given in intuition, i.e. that the conceptual make up of the numerals is primitive. At the ideal level sequences of symbols are evaluated as being of a specific kind. Even real mathematics consists in finitary reasoning operating on forms. It is just this criterion that is applied to the evaluation of ideal proofs. One way that we might therefore understand ideal mathematics is that finitary reasoning is operating upon meaningless symbols while in real mathematics finitary reasoning is operating upon meaningful symbols. We could then understand the unity of real and ideal mathematics as a function of an abstraction which first takes place at the contentual stage, an abstraction which removes the chemical and physical properties of contentual objects to attain contentual forms. Having discovered the appropriate contentual forms we apply these contentual forms to ideal mathematical formulas to determine if they are proofs. Two points then emerge.

The first point is that, clearly whatever these contentual forms are, they are the hallmark of finitary reasoning. As such, they are tied to contentual objects in some immediate way. Detlefsen must therefore argue that the reliability of finitary reasoning is not due to the objects upon which finitary reasoning operates, since at the ideal level the objects of manipulation are purely syntactic elements. Thus, at the contentual level, finitary reasoning and its purported reliability must already be a function of reasoning on forms, on syntactic elements of a specific kind. But this just means that finitary reasoning is already a function of the formal properties
of a formula. That these formulas are meaningful therefore plays no important role in our assessment of a contentual proof as being reliable.

The second point is that, even if Detlefsen could provide an adequate account of how finitary reasoning applies to contentual and ideal mathematics, we would still be left with having to provide a non-formal account (that is, an account that does not take form as a sufficient condition for the content of an ideal formula) of ideal mathematics. Ideal formulas are understood to consist of syntactic elements. A syntactic element is a concrete physical object. The metalevel evaluation of this "object" involves associating a "form" with a concrete object. If our evaluation is to be adequate for ideal mathematics, this "assignment" cannot be given via a formalist account. Implicitly we must be able to assign some meta level regularity to the symbols at the ideal level. Detlefsen intends this to mean that we notice some formal property of the set of formulas being examined. As Charles Parsons has argued, this entails that some non-formal function be entertained to explain how the different relata are related in mathematics. (Parsons, 1990, p. 304) The point is that at the metalevel "assigning" a "form" to a set of syntactic elements is not syntactically characterizable. The objection is that at the metalevel our grasping of a set of syntactic elements as being of a specific form involves a primitive intuition. Yet this is unavailable to Detlefsen. For Detlefsen, even at the contentual level, appealing to the tokens in order to guarantee reliability is only apparent. More importantly, if our metalevel evaluation dictates that concrete tokens are the objects on which our evaluation operates,
and if the objects of contentual mathematics are relevant to finitary reasoning, then any reason that ideal mathematics is different from real mathematics must lie in their different ontologies. The less that is said about how the metalevel evaluates ideal formulas, the more it appears that our metalevel evaluation presupposes the very intuition that Detlefsen must underplay in the case of contentual mathematics.

From the formalist perspective the kinds of intuition necessary are minimal. Fraenkel et al have identified the following ingredients of a primitive kind of intuition that are necessary for the formalist project to get underway:

1. Only the kind and order of the symbols are to be relevant.
2. The meanings of the terms are not relevant, the intended meanings of the extra-logical symbols should not enter into the derivation (unless explicitly mentioned in the axioms).
3. Only a minimal intuition is needed, the ability to determine when two symbol tokens are to be understood as two different symbols or when they are occurrences of the same symbol type. (Fraenkel et al, 1958, p. 276)

Fraenkel et al go on to say of the intuition involved that it is a kind of intuition that requires no intellectual powers at all and can be built into suitably constructed machines. (Fraenkel et al, 1973, p. 276)

The formal system’s freedom from any intended purpose for which the symbols can be interpreted is directly related to their usefulness. (Körner, 1960, pp. 85f.)
The mathematical formalist envisions a specific kind of manipulator. This manipulator is characterized by the minimal intuition that he must contribute to the mathematical enterprise. A kind of mathematical neutrality is developed in which only the person who purges all content, and with it one’s prejudices, can assess the truth of a mathematical formula.\(^8\) Kenneth Manders has referred to this neutrality as endowing the cognitive agent with "conceptual omnipotence". (Manders, 1985, p. 201) Frege in a similar fashion is critical of the assumption that "concepts" are free-floating and independently existing objects. While Detlefsen does not explain all that is involved in the metalevel he must underplay Hilbert’s own view as stated in "On the Infinite". Hilbert, in listing the preconditions, argues as follows:

As a further precondition for using logical deduction and carrying out logical operations, something must be given in conception, viz., certain extralogical concrete objects which are intuited as directly experienced prior to all thinking. For logical deduction to be certain, we must be able to see every aspect of these objects, and their properties, differences, sequences, and contiguities must be given, together with the objects themselves, as something which cannot be reduced to something else and which requires no reduction. This is the basic philosophy which I find necessary, not just for mathematics, but for all scientific thinking, understanding, and communicating. The subject matter of mathematics is, in accordance

---

\(^8\) As will be argued, the universality of the language presents a serious problem for the formalist. For example Kenneth Manders argues that, 

*if every thing is expressible in a language, no special properties of things expressed follow from mere expressibility.* (Manders, 1987, p. 200)
with this theory, the concrete symbols themselves whose structure is immediately clear and recognizable. (Hilbert, 1926, p. 192)\textsuperscript{9}

Kitcher identifies this "intuition" as Kantian. Each recognition of some specific concrete symbol as having a specific structure is given in perception. Implicitly, some version of the "third man argument" is avoided by the fact that "structure" (form) and concrete symbol are not separated in Kantian intuition. At the metalevel this unity is not present. Our assigning a structure to some instrument is not a function of intuition; rather it is a function of its utility.

Detlefsen approaches the formalist idea of what the backdrop for the mathematical enterprise is. One might object that the backdrop is irrelevant to the mathematical enterprise. Clearly the mathematician working in some specific field need never be concerned with giving an account of this backdrop. None the less, two important considerations are relevant. From a philosophical perspective it is surely desirable that our account of mathematics be self-contained. Thus if one is a formalist, the hope is to account for mathematics and any relevant procedures in terms of formalist criteria. If this is impossible, the danger exists that we are unwittingly importing illegitimate procedures into even that part of mathematics in which our formalism seems most secure. The second consideration is from Kenneth Manders. The introduction of extensional elements needs to be justified. A natural way of understanding the metalanguage is to see it as a consistent extension of our object language. Since for the formalist the background language

\textsuperscript{9} This claim of Hilbert's stands in direct opposition to Frege's view that "concepts" are not primitive.
is just the metalanguage, one might ask what the justification for this metalanguage is. The two options that have been discussed are

(1) to devalue the role of the concrete symbols in the object language, and

(2) to understand the metalanguage evaluation as taking place on concrete symbols.

Both accounts purport to provide the needed unity to real and ideal mathematics. My contention thus far has been that neither is successful.

Detlefsen notes that, for Frege, the conclusion is causally effected by the "preceding noetic acts" that the premises and inferences generate. (Detlefsen, 1986, p. 10) Hilbert, like Frege, believes that the correct paradigm for any mathematical proof is a contentual proof. Detlefsen then characterizes the problem as one in which the epistemic potency of the conclusion is the crucial element. Central to the attack on Frege is the concern that Detlefsen mentions regarding the strength of the evidence for ideal proofs.

The Dilution Problem, as we shall present it, shall be primarily concerned with the comparative strength of contentual mathematical proof and its proposed metamathematical replacement. If, on the whole, the proposed metamathematical replacements yield weaker evidence for the propositions proven than do the contentual mathematical proofs which they are to replace, then we say that the proposed replacement produces epistemic dilution. (Detlefsen, 1986, p. 17)

As he argues, this evidence with respect to real proofs is clearly very strong. Detlefsen then continues, arguing that the replacement strategy of the Hilbertian does not imply that the metamathematical replacement be finitary. Nevertheless
it does make this an attractive strategy. (Detlefsen, 1986, pp. 17f.) Thus the Hilbertian must characterize finitary proofs so as to accommodate ideal proofs. Without this accommodation the expressed strength of finitary proofs will always put into question the epistemic potency of ideal proofs. This accommodation is performed by noticing that finitary proofs can be separated into two groups: those which are problematic and those which are unproblematic. The unproblematic proofs are those which are so short and efficient that their conclusions are immediately apparent. The problematic proofs are those which, although finite, are very long and complicated. In these cases the epistemic potency of the conclusion is diminished due to the complexity and length of the proof.

The division of finitary proofs into problematic and unproblematic allows the Hilbertian to characterize the nature of finitary proofs in terms of "the epistemic potency" of a proof. If problematic proofs are finitary and the epistemic potency of a proof remains a function of the complexity of a proof, then "finitary" cannot be characterized by an appeal to intuitions concerning the objects being manipulated. In the case of the problematic proofs, complexity effects our confidence in a proof. It thus provides the Hilbertian the opportunity to characterize finitary proofs as the form of the proof, even in the case of the unproblematic finitary proofs.

One difficulty is in relating this stated view to that of Hilbert. Recall, any appeal to a representation of the numerals is already making an appeal to non-finitary reasoning. Even then, the representation of a numeral by the letter "a",
while already making an appeal to non-finitary reasoning, is informed by the numerals themselves. The form of each sign-token, as Detlefsen interprets Hilbert, is independent of its time, place and chemical composition.

Thus, the objects of finitary thought are sign-tokens, but sign-tokens treated as having only a "form" or "shape". (Detlefsen, 1986, p. 51)

From this Detlefsen concludes that the

scope of a finitary operation is, therefore, never just one or another particular token, but rather the whole class of tokens sharing a given form. (Detlefsen, 1986, p. 51)

The key distinction that arises is that the intermediary steps of an ideal proof are meaningless, and these are evaluated at the metalevel by paying attention to the form of the proof, while real proofs involve their contents directly. It turns out that our intuitions are informed by the "form" of the numerals, the form which constitutes the criterion of whether or not a specific sign-token is a member of some specific, reliable class. Returning to the black box, its internal configurations are to be considered to be appropriately finitary if the sign-tokens it manipulates are of the appropriate form. In the two conditions (1) and (2) above\(^\text{10}\), the explicit mention of syntactic character is meant to differentiate the contentual formulas from the ideal formulas. Applying the two conditions we must conclude that finitary proofs are finitary because they share a form; it is implicit that this form does not merely rely upon syntactical properties. If it only depended upon syntax,

\(^{10}\) These two conditions were described as (1) that the ideal proof is an object with a specific allowable syntactic character, and (2) that objects with this syntactic character have a terminal formula which expresses a true real proposition.
then the symbols could be understood to be meaningless. Characterizing this form as a non-syntactic something which is syntactically exhibited by the numerals undermines any attempt to characterize the syntactic character of an ideal proof and its potency as relevant to the intuitive certainty of contentual mathematics.

Detlefsen’s characterization of Hilbert’s instrumentalism involves two important moves which have already been discussed. The first is the move to find some common characterization of both ideal mathematics and contentual mathematics in terms of the "epistemic potency" of a proof. This move is performed by the analysis of the problematic finitary proofs in terms of "surveyability". The second move involves an analysis of the "objects" of finitary mathematics. Both analyses involve a reassessment of the strength necessary for a finitary proof to guarantee its epistemic potency.

Crispin Wright rightly points out that the Platonist is obligated to explain how it is that we can understand a mathematical statement to be true when that statement goes beyond any possibility of a proof. In such circumstances, can we have a grasp of its truth value? (Wright, 1982, p. 208) Wright argues that the intuitionist is obligated to explain how a statement whose computability is possible, but beyond actual practice, can be understood to have a determinate value. The distinction introduced is reminiscent of Detlefsen’s distinction between problematic and unproblematic proofs, problematic proofs are computable only in principle. His description of problematic proofs is ultimately dependent upon the characterization of capacities and descriptions of the actual world. Thus, if an
idealized person is able to compute at very fast speeds and keep the computations in mind, then the problematic computations become unproblematic. Even then, though, it seems that this idealized person cannot perform an actual computational proof of an ideal statement. Clearly such augmentations of this idealized person’s abilities ignore both what would seem to be mechanically beyond their actual capability and beyond what is nomically possible. Nonetheless, one might wonder, if this person can be conceived to transcend his or her physical and mechanical abilities what would prevent us from augmenting these abilities with additional capacities so that computational proofs of even ideal statements could be given. Any attempt to extend one’s computational skills would, it seems, threaten to effect our concept of the finite. After all, "computation" seems to characterize the appropriate set of finite numbers.

The concept of "being finite" is not generally a relative concept. If it depends upon a description of our capacities and a description of the world, then it may become a relative concept. Furthermore, this person would be forced to perform an infinite number of concrete manipulations, of some set of objects, in a finite period of time. Whether this is only an apparent contradiction will depend upon whether one sees the idealized person’s extended capacities as themselves contradictory, namely as computations which are faster than the speed of light.

It appears that the augmentation of the idealized person’s abilities, so that he or she is able to perform ideal computations, requires that he or she exceed finitistic possibilities. The difficulty is in providing an explanation for "finitistically
possible" so that the definition is not circular. (Wright, 1982, pp. 224ff.) The idea that only certain kinds of computations can be considered as finitistic, presupposes that "being of finite length" is characterizable independently of what some human can actually perform. As Wright points out, even such concepts as 1-1 mappings, or the subset relation, are at root dependent upon the idealization of a person's capabilities.

Let us assume for the time being that "surveyability" can be adequately defined. Furthermore, let us assume that the ability of some idealized human to survey a computation adequately characterizes finitism. This idealized person's ability to perform any finitistic computation requires an appeal to the possible manipulations of concrete symbols that can be performed and kept in mind. Appealing to an idealized person's experiences already outstrips any actual person's experiences. Thus any argument which denies this person's abilities from being augmented so that he or she can perform computations which are faster than the speed of light, on the grounds that this new ability is contrary to actual experiences, will be inadequate. The idealized person already surpasses any actual abilities.11 "Actual experience" is a reference to what we can expect the average person in a normal environment to in fact experience. As van Fraassen points out, "experience is a success term". (van Fraassen, 1990, p. 9) The point here is that under specified conditions a person is expected to perform the task in that

---
11 Appealing to the description of the world, i.e. to the limiting laws of nature, appears to place mathematical truths in the domain of science.
situation. Not only is "experience" a success term, it also characterizes when we expect a person to conform to the norm. The difficulty arises in being able to determine which experiences are excluded and then accounting for those experiences which are outside the norm as being, not merely irrelevant, but excluded as experiences. The implication is that, should our expectations change with respect to our computational abilities, the relevant conditions for which our experience becomes the arbitrator would similarly change what we should accept as being reliable. The relationship between "repeated successful practices" and the incorporation of these actions into the domain of the primitive facts of a field is here acknowledged. Any action which over a long period of time proves to be reliable provides sufficient motivation for an explanation that is acceptable in terms of accepted scientific language\textsuperscript{12}. A case in point is the importing of "new entities" by the creation of new notation within mathematics, e.g. for the imaginary and transfinite numbers. The point here is that the idea of "finitistic computation" is in part dependent upon an idealized person's capabilities and what he or she can do under some description of his or her idealized capabilities. The repeated successful practise of an action may be sufficient grounds for the action in question to be incorporated into the idealization.

As Wright points out

"the conditions under which failure indicates inability in the relevant respect are not sharply determined". (Wright, 1982, p. 232)

\textsuperscript{12} Van Fraassen makes this point in his [m1990].
Wright is concerned with the characterization of the finite by actual 1-1 mappings. The problem he recognizes is that in order to determine that some failure in mapping is due to a relevant inability, we must have at our disposal the set of finite computations. It is this set which allows us to determine that failure is a function of an inability and not a function of a human error.

As Detlefsen recognizes, the Dilution Problem poses a serious threat to Hilbert’s program. If the move to ideal mathematics and the resultant characterizarion of mathematics weakens the notion of mathematical proof, then the net gain of the epistemic potency of a proof may be so diminished so as to negate its worth. Detlefsen’s move is to characterize epistemic potency in terms of reliability. This characterization depends upon the correct characterization of both problematic and unproblematic proofs. The important factor is that if my appeal to van Fraassen’s and Wright’s arguments are correct, then the reliability of a proof is moved outside of the domain of mathematics. More importantly, the set of finite numbers will then be seen to involve a normative concept, but one with sociological criteria and relative to the definition of experience. While this definition may in fact perform a useful function for ideal proofs, it severely weakens our confidence in finitary proofs. Since the importance of finitary proofs results from our confidence in them, any relativity would seriously undermine this confidence.
CHAPTER TWO

Formalism, as I have begun to describe it in Chapter One, entails both a kind of neutrality and a kind of "conceptual omnipotence" on the part of the cognitive agent. Neutrality is placed in doubt when one considers the minimal intuition necessary to generate the concept "iteration of stroke expressions". Conceptual omnipotence is placed in doubt by the fact that the cognitive agents do not have at their disposal "all the concepts expressible in a formal language". (Manders, 1987, p. 201) Frege's attack on Hilbert takes Hilbert's account as incorporating formalist strategies. Frege's arguments are directly related to his attack on Aristotelian logic, that is, on the prioricity of concepts. Contrary to the formalist account of logic, logic as understood by Frege has its own subject matter and it is in light of this subject matter that reliability is to be judged. There is a tendency to understand Frege's logic as the first modern expression of a subjectless notation. Kitcher understands Frege's indebtedness to Kant as consisting in the adoption of Kantian intuition and, thus, of Kant's epistemological theory of what constitutes our "knowing that x". (Kitcher, 1979, pp. 248ff.) At the root of this epistemology is the separation of "form" and "meaning". Logical truths are those truths which are unaffected by meanings. One issue that remains unexplained by this account is Frege's vehement attack on formalist accounts of mathematics. After all, the natural extension of separating form from meaning would be that logic is the study of forms. Frege's attitude towards his own notation, that is, that
it is a new language which does not contain the ambiguities of a natural language (e.g. vacuous terms) and that natural languages should be abandoned, might appear to justify the exercising of meaning from logic. However, I will take this justification as misrepresenting Frege's intentions for introducing his logical notation. Defense of my position will consist of examining Frege's critical comments on Hilbert's geometry and Frege's insistence that his logic necessitates conceptual analysis.

Two different approaches to logical notation can be described as follows. The first is the account rehearsed in the preceding paragraph where "form" becomes the predominant concern of logic. In this version the introduction of an unambiguous and precise language is understood as replacing both the natural language and the conceptual framework this language contributes. This view leads naturally to the view that the language itself is to be studied. The second is an account that takes the "new" unambiguous and precise language as being a function of the conceptual make-up of our natural languages. This new language becomes an ideal expression of the natural languages' content and thus demands conceptual analysis. Conceptual analysis is needed in order to determine the adequacy of the notation for capturing some specific content. It is this second version that I shall assume and only support with indirect arguments.

At the centre of Frege's attack on Hilbert is Hilbert's formalism.¹ The

¹ A contrary view is given by Hintikka who argues that much of the confusion in our understanding of Hilbert's view is due to incorrectly attributing formalism to him. (Hintikka, 1977, sec. 5) Regardless of the truth of this claim, Frege clearly
difficulty in assessing Frege's arguments is demonstrated in the controversy over who Frege's predecessors were, and his relation to his successors. The determination of Frege's arguments as good or bad demands that we understand at whom these arguments were directed. We must also understand the assumptions that Frege shared with his contemporaries and where he parted company with them.

Dummett, in assessing Frege's contribution, places Frege at the birth of modern logic. (Dummett, 1973) Frege's indebtedness to his predecessors is seen as minimal. His writings provide a new start for logic and witness the end of a period of time in which the field of logic was without any notable contributions. Frege's opponents are seen as the Idealists, and their descriptions of reality are seen as assessments of the internal and necessary relations between concepts. On this view Frege's successors are the model theoretic mathematicians. Resnik similarly recognizes Frege as having provided the necessary groundwork for model theory. (Resnik, 1973)

In its separation of content and form model theory recognizes two distinct categories. On the one hand, formal symbols are devoid of semantic content. The only relevant characterizations are those which express pure relations between the symbols. On the other hand, these symbols are taken to have semantic import (content) when some part of reality attests to their usefulness (i.e. truth). The first of these two categories is at least favourable to the idealist's goals. The difference understood Hilbert's work as embodying formalist strategies.
is that for the model theorist, the study of a theory becomes the study of its models. Model theory, and its underdefined conditions for identity, relegates identity to conditions on isomorphic models. For any two models in which all the relations expressed by the theory are realized, the models are seen to be identical with respect to the claims of the theory.

Webb, commenting on Hilbert, argues that

Hilbert's claim is thus that any theory can only describe its basic domain up to isomorphism: its axioms can never distinguish isomorphic models. To repeat the words of Weyl, "the idea of isomorphism demarcates the self-evident insurmountable boundary of cognition". (Webb, 1980, p. 81)

Webb argues that Hilbert’s formalization of geometry is not completely modern and that it rests upon a visual intuition. The modern view rests on "dimensionally neutral ideas". (Webb, 1980, p. 88) The concept of "isomorphism" plays the role of delimiting the boundary of cognition by constraining the criteria that allow us to distinguish two theories. In the case of a holistic theory of meaning the formalism becomes meaningful in the interplay between the axioms of the theory and its models. To this extent any important "intuitions" are formalized and thus no other criteria are relevant. The demand is that our formalism mimic these intuitions in just the right way, such that any two theories which share all the same models are logically equivalent.

According to Webb, Hilbert's "formality" in geometry finds its roots in Hertz’s mechanics of the 19th century. (Webb, 1980, p. 78) Boole had urged that
every system of interpretation which does not affect the truth of the relations supposed, is equally admissible. (Quoted in Webb, 1980, p. 79)

As Webb points out, the move is to a model theoretic account. Implicitly, the formal language in which the axioms are couched play a minimal role. In fact, it is just because of the minimal role of the formal language that the concept of "isomorphism" can play a significant part. The very concept of "isomorphism" assumes that for any description there is a necessary gap between the world and the language we use to describe it. Since there is no mediating factor, there is no unique description of the world. Every description fails in that its characterization is neither unique nor complete. Thus at best we can characterize some situation such that, for any two descriptions (theories), if they describe the very same situations (models), then we cannot demand anything more. It is because "truth" is relative to a model and not to the world that isomorphism plays a dominant role.

Dummett points out that

To have a Fregean grasp of sense, we must have a conception of what it is for a statement to be true, independently of our means of recognising its truth. There appears, however, to be no non-circular way of explaining what it is to have such a conception, or hence of giving an account of understanding that does not presuppose what it purports to explain. (Dummett, 1991a, p. 16)

He goes on to claim that the verificationist strategy is to demand that we replace the notion of knowing what it is for a statement to be true by that of knowing what would rightly lead us to recognise it as true. (Dummett, 1991a, p. 16)

Dummett continues, arguing that the verificationist substitution reflects Frege’s commitments to
meaning as a communally recognised feature of expressions, and understanding as the grasp of that feature. (Dummett, 1991a, p. 16)

At first it might appear that the isomorphic condition is especially adapted to the "communally recognized features" that Dummett is discussing. This relationship is only superficial though. The mere fact that a formal language only involves a minimal content, and thus contributes nothing of significance, already presupposes that there is a neutral starting place. Furthermore, the fact that there can be two theories which cannot be differentiated because of an inherent gap between the world and our language undermines the communal determination of the truth of a statement. By implication, the role that the world plays as a final arbitrator is usurped by the role that the models play. Any account that presupposes communal determination to be absolute threatens to disallow any comparison of communal practices. If the contribution of the uninterpreted theory is minimal and the individual models provide the needed semantic content, then the comparison of models from the perspective of the uninterpreted theory seems illegitimate.

In contrast to Dummett and Resnik, Sluga argues that Frege's opponents are not the idealists. (Sluga, 1980, pp. 13ff.) Frege, in accepting Kant's a priori/a posteriori and analytic/synthetic distinctions, in fact aligns himself with idealist arguments of his time, and with a movement which had been on the decline since the 1830s. Frege's opponents, as Sluga argues, are the scientific naturalists. Sluga describes "scientific naturalism" as consisting of two important components: materialism and psychologism. As the term "scientific" suggests, logic was taken to be comprised of concrete manipulations of concrete symbols; and as the term
"naturalist" suggests, the psychological process from "true belief" to "knowledge" is taken to be the determining factor in assessing the criteria for knowledge. "Materialism" reflects the growing dependence of philosophy on science. By the early 1800s, physiology as a science was taken as providing a viable explanation of how the external world affected knowledge. Increasingly, psychologism posited sensual experience as being on a par with the objective world. As Sluga points out, once sensory experience is understood to be of equal value to that of the external world, there is a strong temptation to see the objective world as constructed out of these sensory experiences and the external world becomes redundant. It is in light of this slipping into phenomenalism that Sluga understands Frege's attack on psychologism. Frege's logical objects are to be understood as undermining phenomenalism without appealing to materialistic criteria. Since Frege's "logical objects" are conceptually generated, the appeal to sensory experience is not necessary. It is interesting to note that Frege's conceptually generated "logical objects" have their origin in von Staudt's geometry of the early 1800s. Wilson points out that Frege's examples of logical objects are at times exactly those of von Staudt. (Wilson, 1992, pp. 162ff.) Von Staudt had introduced the notion of "concept-objects" as a means of justifying the introduction of "ideal points" in geometry. Frege developed the idea of a "concept-object" to include extensions and ranges of values.² "Logical objects" (concept objects) are understood as

² Following Wilson, "Concept-objects" is here used "to represent whatever object it is that we plan to extract logically from an originating concept. Frege, in his practice, invariably utilizes the concepts' extensions - that is, sets - as the
justifying the introduction of extensional elements when the original field and its conceptual framework demand that these objects are necessary. Thus, as Frege argued against the Booleans, his logic demanded conceptual analyses and not merely computation. "Reliability" in mathematics was not a sufficient condition for the mathematical enterprise, since the introduction of extensional elements is not justified without the analyses of the conceptual framework in which we are working.³

Boos distinguishes formalism from radical formalism by the role that an intended model plays. (Boos, 1985, p. 150) While formalism (unlike radical formalism) may have a mathematical domain in mind which the formalism is supposed to capture, this domain is not regarded as being relevant. The supposition is that, while our formalism is embedded in a wider language in which conceptual content plays an important role, once constructed, the formalism is to be free from the wider language’s constraints. Hilbert’s reply to Frege, that the scope of a formal theory’s applicability is unrestricted and allows for new and unexpected applications, reflects the lack of contributing factors that the intended model provides once the formalization has taken place. But even here, Hilbert expresses some qualification; applicability must be guided by the mathematician’s "good will". Boos sees the "good will" qualification by Hilbert as being reflected selected concept-objects". (Wilson, 1992, p. 162)

³ It seems there is a strong similarity between Wilson’s account of Frege and Detlefsen’s account of the "architecture of a field" in Pointcaré’s work. (Detlefsen, m1992)
in the possible interpretations of the formal theory, and that a correct understanding of Hilbert’s argument demands a background "meta-set-theory" which can quantify over these interpretations. The wider language in which our formalism is imbedded is thus the "meta-set-theory". The important point here is that the meta-set-theory is just more formal logic. (Boos, 1985, p. 153) But how are we to understand this theory? Boos argues that Frege’s problem is that Frege is unable to envision a language in which the metalevel theory is not a universally applicable language. The idea is that the paradoxes of set theory are created when language is not sufficiently stratified to prevent quantification over "all there is". Let us assume that the stratification works for the metalevel. Then by induction the assumption is that it works for the nth metalevel also. These increasingly more powerful languages are still just more formal logic, though. Without some explicit understanding of these languages, their usefulness is diminished. What are the properties of these languages? What is a property-of-a-property-(n times)-of-a-property? Does Skolem’s theorem enter in a new guise? Any concepts that this language uses will be extremely abstract and tenuous. How are these concepts able to throw meaningful light onto our object language? It is, after all, the object language about which we wanted to gain knowledge. In fact, Field questions the possible counter-intuitive understanding of "implication" and "consistency" that these higher logics give. (Field, m1991, p. 5)

Putnam is also critical of the idea that one can characterize the concept of "true in L₁" without appealing to the intuitive notion of "truth". (Putnam, 1988, pp.
His criticism rests upon the problem of translation, specifically the translation of the object language in the metalanguage. Borrowing an example from Carnap, Putnam explains the difficulty in the following manner:

Let $L_1$ be a language with just two sentences
Der Mond ist blau (meaning the moon is blue),
and
Schnee ist weiss (meaning snow is white). (Putnam, 1988, p. 62)

We can then define "$S$ is true in $L_1$" as

(1) $S$ is true in $L_1$ iff $(S = "\text{Der Mond ist blau}"$ and the moon is blue),

or,

$(S = "\text{Schnee ist weiss}"$ and snow is white).

(Putnam, 1988, p. 62)

Two important points need to be made. First, the list on the right hand side exhausts the possible substitution instances for $S$ in $L_1$. Second the phrase, "true in $L_1$" only apparently results in circularity. The hope is to characterize the intuitive meaning of "truth" without running afoul of the paradoxes. Thus (1) is to be understood as shorthand for

(2) $S$ is true in $L_1$ iff $[(S$ is spelled D-e-r-space-M-o-n-d-space-i-s-t-space-b-l-a-u and the moon is blue)

or,

$(S$ is spelled S-c-h-n-e-e-space-i-s-t-space-w-e-i-s-s and snow is white)]

and

no other inscription with any other spelling is a well-formed formula of $L_1$. 

46
The right hand side avoids any reference to "truth". The hope is to characterize "truth" in terms of a list which syntactically exhausts all the possibilities without using the problematic intuitive concept as a guide. Thus (1) becomes

\[
\begin{align*}
&\text{[\(S\) is spelled "Der Mond ist blau" and the moon is blue]} \quad \text{iff} \quad \text{[\(S\) is spelled "Der Mond ist blau" and the moon is blue]} \\
&\text{or} \\
&\text{[\(S\) is spelled "Schnee ist weiss" and snow is white]} \\
&\text{or} \\
&\text{and} \\
&\text{no other inscription with any other spelling is well-formed.}
\end{align*}
\]

(Putnam, 1988, p. 64)

The list that originally appeared on the right hand side of (1) now appears on the left hand side. As Putnam points out, our new definition is a tautology. More importantly,

every language which satisfies the syntactic restriction satisfies this"!
(Putnam, 1988, p. 65)

Assume that another language, \(L_2\), is given by the following two sentences:

"Der Mond ist blau" is true iff "the sky is blue",

and

"Schnee ist weiss" is true iff "water is white".

We then want to show that (1) excludes \(L_2\) as a correct analysis of "truth" in German, but as Putnam points out, \(L_2\) is not excluded and any language which
meets the syntactic criteria is adequate. The German "meanings" of the two phrases

   der Mond ist blau,

and

   Schnee ist weiss

which appear on both sides of the biconditional after the syntactic criteria (i.e. the spelling) rely on our intuitive understanding of what a correct translation is. That is, they rely on our intuitive understanding that the phrase "der Mond ist blau" is not translated to mean "the sky is blue". The concept "correct translation" is supposed to imply that one can give an account of "the language" one is examining by listing the syntactic criteria and that only these conditions need to be met in order for the translation to be correct. The argument that one should not change the meaning of the terms relies on an intuitive understanding of what these meanings are, and on the inadequacy of the syntactic characterization of capturing what a "correct translation" is. Putnam's argument relies upon our accepting the view that syntactic characterization always presupposes a certain background language which we hold fixed, and that without this background language our "translation" becomes an empty condition. (Putnam, 1988, pp. 64f)

Carnap's "semantic rules" perform the function of identifying $L_1$ as a fragment of German. Putnam argues that these rules must be given in the metalanguage and thus that they must conform to the "syntactic conditions" of the metalevel. The difficulty arises when one attempts to give some meta-
metalevel rules for these rules. At some level the syntactic characterization will need to rely on our intuitive understanding of the concepts being employed in order for us to account for the correctness of the "translation".

The request is thus that the "translation" exhibit a certain amount of "good will" by the translator. The question that arises for the Hilbertian is this: Under what conditions does this "good will" exceed the very conditions to which the Hilbertian is entitled? Mark Wilson, in examining Frege’s position on Geometry, provides a contemporary setting for the problem at hand. (Wilson, 1992) Wilson is concerned with the question of how a solution to a specific problem in mathematics can be justified when the solution is given in an extension of the domain in which the problem was first expressed. What justifies one extension over another? Without some justification, these extensions will appear to be arbitrary. As a means to an end, any calculatory procedure is surely as good as any other. Frege’s attack on the Booleans addresses the question of the status of the Boolean notation. It raises the question of whether their notation represents a universal language, i.e. a language in which the content is clearly expressed in the notation, or whether the language merely provides a mechanical means of attaining a solution. (Sluga, 1987, p. 81f.) Frege, in his attack on Boole, is appealing to the role that notation plays and the subject matter it is meant to capture. Contemporary attempts at providing an answer to the question of how to justify a solution in a domain different from the original domain rely upon the "intuitions"
of the mathematicians and the availability of models that might conceivably satisfy the uninterpreted theory.

The "official" position, dominant since the start of this century, maintains that any self-consistent domain is equally worthy of mathematical investigation; preference for a given domain is justified only by aesthetic considerations, personal whim or its potential physical applications. (Wilson, 1992, p. 152)

Hilbert's claim that with "good will" our theory's scope of application may surprise us is clearly entrenched within this explanation. As Wilson points out, any attempt to solve this problem and avoid the contention that the solution constitutes a change of topic and is thus not a relevant solution, relies upon identifying the "optimal setting" in which our notation and theory lives. Yet this clearly demands that we take seriously the "meaning" of the concepts involved. (Wilson, 1992, p. 152) According to Wilson this depends "solely upon semantic concerns". (Wilson, 1992, p. 150) Frege's attack on Hilbert's independence proof addresses this problem.

Frege, in his criticism of Hilbert's independence proof, questions the idea that a background language can contribute the necessary content the formalist needs. Wilson draws attention to Frege's explanation of how the field of mathematics grows. The attempt is to explain the introduction of "new" mathematical objects (the extension elements) as a function of the nature of the mathematics being studied. The introduction of ideal elements becomes a necessary condition for the explanation of the content of the field in which one is working. Unlike Detlefsen's ideal mathematics, these "new" elements respect the
original content and the conceptual make up of this domain. The relationship of these new elements to the old domain does not need to be justified. Frege’s attack on the independence argument is to be seen in light of this problem.

Hilbert had proposed that the independence of some axiom could be shown in the following manner. Assume that our theory, T, is a formal representation of the axioms of Euclidean geometry, that the sentence A represents the conjunction of axioms 1 to 4 and that B represents axiom 5. T is an uninterpreted theory. If B is the axiom we are attempting to show to be independent from A, then we need to find a "structure" which satisfies A but not B, and another structure which satisfies both A and B.

In a letter to Hilbert, Frege attacks the notion that Hilbert’s argument can show a relevant independence. (Frege, 1903, p. 37) In the first place, it is assumed that for any theory the axioms of the theory provide the implicit definitions used in the theory itself. Thus if we find the letters "p", "o", "i", "n", "t", conjoined in a formula, then in order to determine the meaning of "point" we must first determine the kind of relations that the term "point" enters into with the other terms of the theory. Apart from the relationships that this word enters into, no other constraining factors dictate the kind of referent that these terms can have (i.e. the assignment of names to the objects of the domain).

Frege understands Hilbert’s use of the formula in which the term "point" appears to be functioning as a first level concept. (Frege, 1903, p. 36) Frege’s rationale is that if a specific single point is an object, then it is by quantifying over
these objects that we obtain a first level concept. Frege argues that if we look at Hilbert’s axioms (for Euclidean geometry) we in fact find that the concept point plays a second level role. The complaint is that Hilbert is mixing levels. Frege’s idea is explained using the following example. Let "single points" be the objects of the domain and let A and B be two different models, where A satisfies the axioms for Euclidean geometry and B satisfies the axioms for some specific non-Euclidean geometry. Then both A and B name different objects as "points". In the one case, the A theory allows us to define being an A-point (objects which fall under the first level concept of being a Euclidean point) and in the other case the B theory allows us to define B-points (objects which fall under the first level concept of being a non-Euclidean point). A-points and B-points are first level concepts, they provide the conditions under which A objects satisfy the first level concept of being an A-point and B objects satisfy the first level concept of being a B-point. Frege argues that in order for the two concepts to be compared there must be a concept in which the two concepts, A-points and B-points, fall. (Frege, 1903, p. 37; 1906, p. 102) The important point is just that we are making a claim about the relationship between the two concepts, being an A-point and being a B-point, which requires a second level concept. This view is clearly distinguished from the model theoretic account. There, the uninterpreted theory with its minimal content is understood as conceptually neutral and completely universal.

In his 1903 letter to Hilbert, Frege asked if one could determine whether my pocket watch together with this point determines a straight line. (Frege, 1903, p. 31)
The typical model theoretic answer is Yes, given a favourable mapping of the axiom

Two points determine a straight line

onto a structure (model) in which "this point" and "my pocket watch" are taken as two different points and the other Euclidean axioms are similarly satisfied. Clearly, if the first level concepts of being an $A$-point and of being a $B$-point fall under\(^4\) the second level concept of being a point, and the two different objects named by the two terms "this point" and "my pocket watch" fall under the first level concept, being an $A$-point, then the axiom will be satisfied.\(^5\) Frege’s reconstruction of Hilbert’s independence proof introduces an important role for the background language of a theory. At each concept level a higher ordered concept performs the task of organizing and identifying, from a myriad of properties, just those properties that fall under it. In Frege’s 1884, *The Foundations of Arithmetic*, Frege uses a similar idea to criticize his predecessors’ attempts to formalize arithmetic. Frege’s critical remarks depend upon Hilbert’s failure to provide adequately the conceptual backdrop for the role that the symbols in the independence proof play. Frege points out that Hilbert’s axioms must be understood as pseudo-axioms with their corresponding pseudo-propositions. Since

\(^4\) Contrary to Frege’s own use, I will use "fall under" for both expressing the relationship between objects and first-level concepts and the relationship between first-level concepts and second-level concepts.

\(^5\) See Frege 1906b, pp. 36f. Frege’s criticism of Hilbert’s independence proof will be taken up again later in this chapter.
the symbols in Hilbert’s axioms do not have determinate referents, the pseudo-propositions do not present determinate thoughts. (Frege, 1906b, pp. 82f.)

In the "Foundations" Frege spent the first third of the book attempting to show that the previous attempts to define the "numbers" had failed and that this failure was due to a bootstrapping operation. Previous attempts had begun with primitive objects, either mental or physical, as given in perception or apperception. The concept "number" was then constructed out of these specific numbers. Repeatedly, he shows that these constructions are not adequate. In the "Foundations", his concern is with the ability to delineate the characteristics of "number" apart from the myriad of properties that "objects" have. Frege defines a "characteristic" in the following manner:

A characteristic of a concept is a property an object must have if it is to fall under that concept. (Frege, 1903, p. 35)

Frege shows that any attempt to get the individual numbers from objects or their relations to objects is misguided, in that the object’s diversity or identity does not allow the requisite abstraction to get underway.

The concept has a power of collecting together far superior to the unifying power of synthetic apperception. (Frege, 1959, p. 61)

In the example that Frege gives immediately after this quote, he states that it would be impossible to join the inhabitants of Germany together into a whole in order to count them. Without already knowing what the conditions are for deciding when something is that kind of thing which we are numbering, we would not know whether or not to include John, who is visiting, as one of the things to be
numbered. The concept of "inhabitant of Germany" brings together all the relevant objects which can then be numbered. Now one might ask, What is a number? Frege provides the "fire risk" analogy as a partial answer. Says Frege we do not acquire this concept by first building a house with a leaky chimney. Instead, once we have the concept of "fire risk" we know what kinds of things to look for. Leaky chimneys adjacent to thatched roofs fall under the concept. When Frege begins to speak about numbers, it becomes apparent that we need a unifying concept in order to unite all the individual numbers. The idea is that in a formula in which symbols are used to represent numbers, we need to assume that these symbols capture the relevant properties of the number concept, but in order to achieve this end we need to know the common part of the set of individual numbers. Frege argues that our symbols are governed by a concept which captures the common part of all the individuals in virtue of our constructing the concept. What is preserved is just this common part. In his attack on the Booleans, Frege argued that his logic was not governed by the mere concern for finding a solution.

Speaking of Frege, Sluga says that

it is in any case fairly clear that he used the Leibnizian sense of calculus when he called Boolean algebra a calculus. For he meant by that, first of all, that it was a device for carrying out calculations .... Boole had shown how one could solve problems of class and propositional logic through mechanical procedures performed on algebraic symbols. (Sluga, 1987, p. 84)

Frege argued that more was required of logic, and that his notation was made possible only through conceptual analysis.
With respect to Hilbert, a radical change occurs from real to ideal mathematics. Ideal mathematical symbols lose all contact with the number concept of which our general laws speak. The gain in content is achieved in the metalevel proofs. In the model theoretic account, our uninterpreted theory and the metalevel language provide the needed content to facilitate a comparison. It is in virtue of these general propositions that our numbers apply, since it is only under some concept that we are able to understand the common part of what is being "counted". In his response to Hilbert, Frege states that definitions, "collect a manifold content into a brief word or sign". (Frege, 1903, p. 23) For Frege, mathematical symbols are introduced by definitions, which capture in a simple fashion just those properties that the concept brings together. The point is that once we have the concept "number" we can explicate characteristics of the numbers, since the concept has already demarcated the "set". The concept abstracts away, from any number just that which we find to be the common part of the numbers. The introduction of a symbol takes this common part and via a definition presents it in an efficient and simple way. But for all this, the symbol is dependent upon the concept, and gains its use because of the concept by which it is defined.

As already mentioned, Frege broke away from Aristotelian logic by introducing quantification. His notation allowed him to explain how judgments affected our understanding of the notation.

The same proposition \( \varphi(a,b) \) can at one time be regarded as a function of an argument \( a \), at another time as a function of an
argument $b$, and at yet another as a function of both $a$ and $b$. And again, if $\varphi(a)$ is generally regarded as a function of $a$, we can also at other times consider it a function with $\varphi$ as argument. (Sluga, 1980, p. 85)

Dummett has argued analogously that in Frege’s philosophy of language "singular terms" and "names" are not independently identifiable categories. It is only within the context of a sentence that terms play the role of being a singular term. Sluga argues that Frege’s "process of analysis" can be explained in terms of the transition from one language to another.

Even if every operation within a given language can be understood as a process of mechanical reasoning, the transition from one language to another cannot be of this kind. It requires insight and understanding. A proposition in the first language will appear to us with a certain structure. Analysis of that proposition involves a translation of it into another language in which that proposition is assigned a new structure. But that move is possible only if our understanding of the proposition is not limited to a grasp of the structure through which the proposition is presented to us, if it does not merely consist in the ability to mechanically manipulate the elements of that structure. The construction of a logical symbolism demands not only calculative reasoning, but also and first of all conceptual, philosophical analysis. (Sluga, 1987, p. 85)

This second claim is especially relevant in Frege’s second version of Hilbert’s independence proof. Unlike Hilbert’s original proof, Frege changes the languages of our two theories, thus preserving a unique referent for each term. "Euclidean parallel lines" are clearly distinguished from "non-Euclidean parallel lines" by the fact that the two terms participate in different languages. Frege then argues that we might compare these two different languages. Frege provides an important caveat to the range of languages that might be used. He argues that the logical constants, i.e. "and" and "or", must stand in a transparent way to the logical
constants of the other language. Frege’s concern is that the "logical constants" capture the same content.

Just as the concept point belongs to geometry, so logic, too, has its own concepts and relations; and it is only in virtue of this that it can have a content. (Frege, 1906b, p. 109)

Logic, in virtue of its content, takes the logical constants as given, and not open to substitution. Thus Frege argues that in an inference we may substitute "king" for "desert" and "Charlemagne" for "Sahara", but we may not substitute the relation "identity" for the relation "lying on a plane". If we take Hilbert’s independence argument to represent a proof in the language of model theory, a problem arises. If no counter example is found (if no model is found which satisfies A but not B) we would be in a position to affirm the logical dependence of the set of axioms A and the axiom B. Our not finding a counter model must here be understood to mean that "no such model can exist". Wright’s comment, that we cannot distinguish failure from impossibility is relevant here. Davis argues that mathematicians are never in a position to affirm an impossibility. He argues that statements of impossibility in mathematics, such as "It is impossible to square the circle" and "It is impossible to set up a one-one correspondence between the set of integers and the set of real numbers" have in fact been converted into "useful and logical possibilities". (Davis, 1987, p. 171) His argument depends upon the fact that mathematical problems are always problems expressed in a specific domain and that domains change. With these changes, impossibilities are converted into "useful and logical possibilities". History has shown that
impossibilities, when structured in a new field (domain) have acceptable answers.
The obvious objection that the original "impossibility" has been redefined in the
new field (i.e. that in the new field the formula "means" something else) demands
that we can provide a strict criterion of what constitutes an acceptable answer⁶.
Frege’s concern can be expressed in the following manner. If $A$ is a Euclidean
model and $B$ is a non-Euclidean model, then the introduction of $B$-points as the kind
of extensional elements that are relevant to the Euclidean axioms must be justified,
in order to show how the fifth postulate is not dependent upon the first four. Any
justification of $B$-points as representing the correct kinds of elements must assume
that $B$-points are points of some kind. That is, we must provide some criterion by
which we can relate the concept of being a $B$-point to the concept of being a
point. For Frege this means that our theory cannot be uninterpreted. The
background language and our uninterpreted theory must assume the very result
that we wish to demonstrate, that the non-Euclidean concept "$B$-point" and the
"objects" which satisfy this concept are related to the concept "$A$-point" in the
relevant way.⁷ Uninterpreted theories do not allow us to determine for some
application that that application is done in the correct spirit. In the case of Hilbert’s
independence proof, the uninterpreted theory is said to be satisfied by one model
but not the other. The satisfaction relation is thus a referent endowing function.

⁶ Similar objections are voiced against the Löwenheim-Skolem theorem, i.e.
that the "membership relation" changes. Manders' excellent 1987 paper is an
attempt to provide criteria for change.

⁷ In Chapter Three the model theoretic argument will be examined.
Thus, in the case of our two different non-isomorphic models, our terms are understood to "mean" different things. Clearly if we find model theory as constituting a reliable test of independence, we must be in a position to delineate how these models constitute a separate class from the new fields Davis is concerned with. The problem is that, since the language of model theory is generally set theory, and set-theory's expressed strength is its universal expressibility, even our new fields will be expressible in our set theoretic language. Manders goes on to express his doubt about the claim of conservativeness. In the case of model theory the objections are pressing. Manders argues that conservativeness may be a necessary reason to justify the use of some extended theory, but that it is not a sufficient one. Manders' primary concern is with how different mathematical structures, and the meaning structures they provide, contribute to our understanding of a mathematical problem. The mere fact that a solution to a problem is found in an extended domain is not sufficient reason for our working in that field. What mathematicians want is a domain in which the structures are made clear. Set theory's universal expressibility confounds the problem of finding a domain which provides this meaning structure. If we take the language of our model theory to be set theory, then the extended domain, whose solutions are conservative over intuitive arithmetic, is set theoretically definable. Our original field, intuitive arithmetic, is also set theoretically definable. Neither

---

8 See Manders, 1985, p. 200, on the problem of the universal expressability of set theory.
theory, as expressed in the language of model theory, will provide criteria that allows us to justify any preference of domains.

Frege argues that holding fixed the "meanings" of the logical constants in the two languages being compared imposes a structure on the inferences allowed. This enables us to compare the two results. For Frege, the characteristics of a concept are brought into relief by an act of judgement. Speaking of physics and gravitation, Frege argues that chemical properties are ignored. The implication is that physics is governed by specific primitive relations, and these relations capture the essential content being studied.

One might argue that a theory is nothing more than a set of well-formed formulas of the language and that conceptual analysis is unimportant. Notice, the background language is clearly set theoretic. Two important points need to be made. In the first place, as Manders points out, from the fact that everything is expressible in a language, nothing of importance follows. In other words a completely formalized language provides no clarification. (Manders, 1985, p. 200) This problem relates back to Manders' contention that modern justification for one domain over another is extremely weak. Secondly, and more importantly, the language of set theory in fact provides a background language. For anyone not familiar with the concept of an "ordered-pair", the explanation of an inference as an ordered-pair is not helpful.

Hallett explicitly states that Hilbert's purpose was in part to reduce semantic problems to syntactic problems, presumably of a language which is complete.
Hallett goes on to argue that Frege's answer is problematic because of the "extra-logical" content that he must introduce, but this is only a special problem if we take the question of the justification of mathematical ontology as a problem divorced from the justification of ontology per se. (Hallett, 1990, pp. 198f.) Admittedly, any attempt to justify mathematical ontology in this fashion must explain how mathematical terms refer. Appealing to the difference between mathematical and scientific terms ignores the conceptual backdrop necessary to generate reference in non-mathematical contexts.

Clearly any comparison of two different languages presupposes a commensurate concept by which these two languages can be compared. As Sluga points out, Frege's break from Aristotelian logic is a criticism of the priority of concepts over judgments. Frege is critical of the presence of concepts prior to the determination of a concept by the act of judgement. (Sluga, 1987, p. 85)

Frege incorporates the mapping of the two languages being compared into the content of the languages themselves. For Frege, the conceptual content is given by the act of judgement, and the act of judgement determines the conceptual make up of the field of inquiry. Thus the relationship of the two different languages being compared is always held fixed, and thus comparable, within a specific range of choices by the fact that the logical constants hold fixed their meanings.

In the Fire Risk case, the concept is already given and provides the requisite partitioning of the properties into characteristics and irrelevant properties. Frege's
concern with Hilbert’s project is that the second level concept is being completed by an incomplete first level concept. Since the first level concept does not obtain a completed reference, our second level concept contains a gap. Clearly this means that there is no concept present sufficient to unify the myriad of properties which characterize the requisite characteristics and none is available from the uninterpreted theory. In his letter to Hilbert prior to his article of 1903, Frege provides the following version of the "pocket watch" argument:

Your system of definitions resembles a system of equations with several unknowns, where the solvability and particularly the univocity of the determination of the unknowns remains doubtful. If the latter did obtain, then it would be better to give these solutions, i.e. to explain each of the expressions "point", "straight line", "between", separately by means of what is already known. I do not know how, given your definitions, I could decide the question of whether my pocket watch is a point. Already the first axiom deals with two points; therefore, if I wanted to know whether it held of my pocket watch, I should first of all have to know of another object that it is a point. But even if I knew this, e.g. of my fountain pen, I should still be unable to decide whether my watch and my fountain pen together determine a straight line, because I should not know what a straight line is. Furthermore, the word "determine" would also occasion difficulties. But even if I were to understand the words "point" and "straight line" as in elementary geometry and were given three points on a straight line, given your explanations and the axioms belonging with them I should nevertheless be unable to decide which of these points lies between the other two. I should not even know what investigations I might have to conduct to this purpose. (Frege, 1903, p. 18)

Frege’s concern here is not merely that Hilbert’s definitions involve only two unknowns. The very idea that something is an "equation" already allows us to determine this "object’s" characteristics. Notice, Frege not only questions the meaning of the term "straight line". He also questions our background language’s
(i.e. our metalanguage's) contribution and the role that the term "determine" is allowed to play. Boos had argued that in an appropriately stratified language the metalanguage provides the needed characterization by lumping together all those first order-models. This is implicitly implied in Boos' argument when he relies on the meta-set theoretic logic in order to characterize the first-order models. For Frege, the fact that this uniquely characterized model is taken to be the "intended model" for Peano arithmetic is completely unexplained in the formalist project. One must already have in mind some model such that the concept "the smallest model" in a second-order logic makes sense. In the case of Hilbert's definitions (axioms) one must similarly already have a concept in mind to make sense of how my "pocket watch" and the "pen" determine a "straight line". Without this requisite conceptual backdrop the axiom is meaningless.

Returning to Hilbert's independence argument, Frege's characterization brings out the requisite conceptual backdrop. A-points and B-points are brought together by the appropriate third level concept "point". Two options are open to Hilbert. The first is that, if we stipulate that the third level concept is not needed (contrary to Frege's opinion), and that our second level concepts provide adequate background content, then our two models A and B play the role of specifying two different theories. In the one, A-points and A-axioms exist, while in the other, B-points and B-axioms exist. Thus, the A-parallel axiom does not "mean" what the B-parallel axiom "means". Frege's question concerns how it is that one is related to the other in such a way that logical independence is demonstrated. Frege's
concern is with the status of the "uninterpreted" theory. The uninterpreted theory provides us with no conceptual content. To make his point, Frege translates one of the axioms into the following gibberish.

Every anej bazet at least two ellah

and then asks

What is anej? What is an ellah? (Frege, 1906b, p. 55)

Thus in the case of the axiom, "Two points determine a straight line", if we fill in the uninterpreted part with the term "A-point" or the term "B-point", then we obtain two different theories whose relationship remains unexplained. In Frege's response to Thomae, a formalist, Frege asks why the symbols of our formalized arithmetic look like the symbols of our non-formal arithmetic. He then replies,

because formal arithmetic cannot do without the sense which its objects have in nonformal arithmetic. (Frege, 1908, p. 137)

The second alternative is to introduce a higher-order logic with the hope that it can group together the appropriate syntactic characterizations. But as Putnam has argued and Frege could reply, these higher-order logics are only more syntax. We are left without any adequate explanation of what the appropriate characterizations are.
CHAPTER THREE

In this chapter, I will show how the arguments which Frege directed against Hilbert can be applied to Wilfred Hodges’ model theoretic explanation of non-logical constants. I will show that Hodges’ account, of how a proper understanding of the non-logical constants can avoid Frege’s claim that only a determinate thought can be true, necessitates a specific conceptual backdrop. Hodges recognizes that Frege’s formalism does not represent contemporary formalist views. Frege’s formal language is a fragment of natural language. Hodges fails to appreciate Frege’s intent in resisting the development of a subjectless notation. In Chapter Two I argued that Frege’s concerns are directed against the supposition that conceptual analysis is not necessary in logic and that logic does not make any special allowance for the background language in which logic is practised. I will argue that Hodges must assume a background language in order for his explanation of a non-logical constant to work and to avoid Frege’s criticisms.

In "Truth in a Structure", 1985/86, Hodges traces the first appearance of the phrase "truth in a structure" to the Tarski and Vaught paper of 1957. The late appearance of the concept of truth-in-a-structure is explained in Hodges’ paper by the fact that earlier writers lacked a clear understanding of the role of the non-logical constants. The earlier languages of Frege and Peano are
fragments of natural language, with the meanings and grammar made more rigorous and with peculiar symbols in place of conventional words. (Hodges, 1985/86, p. 143)

Unlike these new languages of logic and unlike languages sharing their characteristics, the Frege-Peano languages make no allowance for the peculiarities of the non-logical constants. Referring to the latter languages, Hodges continues:

Unlike English, they have no indexical expressions: there are no demonstrative pronouns or verb tenses or persons. So sentences of a Frege-Peano language are generally true or false outright; we shouldn’t expect them to be true on Tuesday but false on Wednesday. (Hodges, 1985/86, p. 143)

The new languages, first-order languages, have four kinds of symbols. These symbols are separated according to whether they are constant or variable expressions. "Constant expressions" are common to both Frege-Peano and first-order languages. In analogy to English, constant expressions do not behave as indexicals and do not need a demonstration, or a time, or a place to fix their referent. The meaning of a constant expression is fixed as a part of the language. If "&" is the symbol for "and", then "&" always has its typical truth functional meaning. The two kinds of symbols that make up the constant expressions are truth-functional connectives, which have a fixed logical meaning, and individual variables (bound or free), which have no meaning. "When [the individual variables] occur free, they mark a place where an object can be named; when bound they are part of the machinery of quantification". (Hodges, 1985/86, p. 144)

Variable expressions are made up of the remaining two kinds of symbols. The first of these is the non-logical constant. This kind includes symbols for the
functions, relations and individual constants of the language. Unlike the constant symbols they have no fixed reference; only in a structure or situation do they have a referent. Thus if \(a_2\) is a name in language \(L\), then depending upon the domain to which the theory \(T\), in the language \(L\) is being applied, \(a_2\) can name different objects. For example, in one domain, \(a_2\) may name a space-time point, while in another domain it may name an integer.\(^1\)

The second kind is the quantifier. Quantifier symbols always have the same meaning.

These always mean 'for all individuals' and 'for some individuals'; but what counts as an individual depends on how we are applying the language. To understand them, we have to supply a domain of quantification. (Hodges, 1985/86, p. 144)

Generally in defining a structure or model we give an n-tuple - usually a 2-tuple or an ordered pair, \(<U^0,R^0>\) - where \(U^0\) is a restricted domain (a universe) and \(R^0\) are the relations that hold in the domain. Quantification is explicitly restricted to \(U^0\) and \(U^0\) demarcates the kinds of things, the elements of \(U^0\), which can be named. Without this demarcation the set of relations, \(R^0\), is ill-defined, since we would not know of what kinds of "things" the relations are said to hold or not hold. Take, for example, the universe, \(U\), such that \(U\) is unrestricted, and

\(^1\) The constant/variable expression distinction is due to Tarski. (Hodges, 1985/86, p. 143) Contemporary logic books, typically divide first-order languages into the logical and non-logical symbols. See Shoenfield, 1967, Chapter One. These two descriptions are not synonymous, e.g. the universal quantifier, in Tarski's language is a variable expression, while in contemporary logic books the universal quantifier is a logical symbol.
let \( R \) be represented by the standard "less than" sign, \(<\). Since our universal quantifier picks up on "everything", we should be able to say whether or not

(1) \( \text{"John} < 3\)"

is true or false, for some specific instantiation of the appropriate universally quantified sentence. Clearly, (1) is nonsensical when "John" refers to a person, "3" refers to the number 3, and "<" has its standard meaning. By implication \(<U^0,R^0>\) delimits the \( U \) (the unrestricted universe) such that we know of what we are speaking. Let \( L^0 \) be the language of some theory \( T^0 \) for which \(<U^0,R^0>\) is a model and let

(2) \( \text{"John} < 3\)"

be a well-formed formula of \( L^0 \). Under these conditions the symbols of \( L^0 \) that make up formula (2) have only minimal meanings or are only minimally interpreted, and thus do not have the standard English referents. It is not merely that if \( s^0 \) is a member of \( R^0 \), then \( s^0 \) is a relation or a function, but that \( s^0 \) represents the kind of "thing" of which it makes sense to say that it can hold of those things which are being named, of the elements of our \( U^0 \). If our language is divided into this kind of order, I will refer to it as having the minimal meaning. Generally this restriction is voiced in terms of the interests of the mathematician. Hilbert in fact explicitly refers to applying a theory tactfully. (Hilbert, 1899, p. 14) The term "minimal meaning" refers to the minimal characterization that a language needs in order to
capture a relevant topic. For Hodges the non-logical constants have only a minimal meaning. The structure provides the rest.²

Two observations can already be made. First, Hodges’ comparison of the Frege-Peano and first-order languages is in part motivated by the range of possible situations³ that could satisfy the formulas of T. To adopt a phrase of John Etchemendy, "models represent each semantically significant configuration of the world" (Etchemendy, 1988b, p. 95) In the case of the formulas in a Frege-Peano language, these formulas are restricted by the semantic components of English (categories and other semantic considerations of the background language). First-order languages, and specifically non-logical constants, are not restricted by these semantical considerations and are thus seen to be more general. Our faith in the proposition that models do in fact represent each semantically significant configuration of the world reflects our faith in first-order conventions as being maximally expressive without making any significant semantic contributions to the formalizations. The degree to which one has doubts about first-order conventions will be reflected in doubts about models being able to represent every semantically significant configuration of the world and, thus, in doubts about the expressive power of our first-order language. One might object that any theory expressed in some language of logic which has a model will be translatable into a first-order

² One can of course take the relations to be a specific subset of the domain. This does not adversely effect the argument as not any subset will do.

³ By "situation" I will always mean "how the world could be", while "models" or "structures" are explicitly mathematical entities.
theory with the same model. In order for this objection to go through one must have a clear conception of what kinds of things models can be, independent of any specific language. Without independent criteria the possibility exists that the set of models and the set of situations are not strictly co-extensive.

Second, the semantic component of first-order languages must be kept to a minimum in order for the structure to play the role it plays. Any substantive conditions placed on the formal language and the formalizations that it allows become restrictions on the kind of models that can be said to satisfy the theory in question. This freedom from semantic considerations plays an important role in Putnam’s attack on formalists in mathematics. Putnam’s argument that no formal theory can characterize, by a formalization in its own language, an intended model, points to the freedom that the formalizations are supposed to enjoy. The inability to characterize some specific set of models as special from within the theory can be seen to be a result of the freedom from semantic commitment of the first-order language.

Wilfred Hodges addresses the question “Why do we need to introduce non-logical constant symbols which are not variables and yet have no fixed reference?” Hodges divides the question into two parts:

(1) "[W]hy can’t we assume that the non-logical constants always have a fixed interpretation?”, and

(2) "[W]hy can’t we treat non-logical constants as free variables?” (Hodges, 1985/86, pp. 145f.)
With respect to (1), Hodges replies that in mathematics we want to be able to compare structures. Take, for example, the theory $T$ (a set of first-order sentences) and some sentence $A$ which is a member of $T$. We want to be able to know that $A$ is true in $M_1$ and $M_2$, where $M_1$ and $M_2$ are the two structures we want to compare. It is because $M_1$ and $M_2$'s similarities and differences can be assessed with respect to the sentence $A$, where $A$ is common to the two structures, that we can give an account of how these two structures compare. Had we been forced to assess the similarities and differences with respect to two sentences say $B$ which is true in $M_1$ and $C$ which is true in $M_2$, the comparison would have been incomplete. We would still have to compare the two sentences $B$ and $C$. Since $A$ is a first-order sentence in which the referent of the non-logical constants is not fixed, we are able to compare the two structures with a common metric. Quoting Tarski, Hodges argues that the non-logical constants do not have a fixed interpretation and thus the results one obtains regarding some theory, say real-closed fields, can be extended so that there is no sentence expressible in our formal system of elementary algebra which would hold in one real-closed field and fail in another. (Hodges, 1985/86, p. 145 quoting Tarski)

It is important that Tarski explicitly refers to "our formal system". It is only in this context that discovering the similarities and differences is relevant. What allows us to move from the comparison of two structures to "all structures" is that the
non-logical constants only involve minimal meanings, and the relevant topic can be characterized in a first-order language.\textsuperscript{4}

With respect to question (2), Hodges provides the following example to explain the difference between the non-logical constants and free variables.

All the Brighton bombers have been arrested. (Hodges, p. 146)

Hodges is concerned with what makes the above sentence true or false at some time or place. If we let $B$ be the class of Brighton bombers and $A$ the class of people who have been arrested, then the sentence is true iff the class $A$ includes the class $B$. As Hodges points out, providing an assignment function such that given the two functions, "$x$ is a Brighton bomber" and "$x$ has been arrested" and then assigning things or people to "$x$" is not relevant to whether the sentence is true. What is important is assessing whether the determined classes $A$ and $B$ are related in the appropriate way, by inclusion. Following Hodges’ example we will understand the \textit{structure} of a time and place to be whatever features of the time and place determine whether certain sentences are true or false. Structures are then those features that make determinate our two classes $A$ and $B$. In group theory the relation symbols used in the axioms for the product operator are identified by some specific structure as being an additive or multiplicative operator.

\textsuperscript{4} An especially interesting example is the Lindenbaum structure. The method involves quantifying over the language and forming equivalence classes under some specific set of descriptions. That is, elements are grouped together and treated as a single element under some description. The formalism cuts up the "domain". It is not clear that without any assumptions about the language that this would be possible.
and in so doing name parts of the structure. (Hodges, 1985/86, p. 146) The assignment of objects to the free variables is then understood as a function of the determination that the structure provides. The failure of accounting for the truth or falsity of a universally quantified sentence by observing the assignment of objects to the variables is apparent when the universe is infinite.

Hodges' account of non-logical constants presupposes a sharp distinction between the theory in which only minimal meanings are involved and the model. The minimal meanings provide the kinds of distinctions that can be assessed and the models provide the context in which these distinctions are to be judged. Hereafter "context" will always refer to the necessary conditions for a term to refer, namely the individuation of the object in a time and at a place. Paraphrasing Hodges' argument thus far one can say that the sentence "All the Brighton bombers have been arrested" is true or false, not in virtue of an assignment of things to "x" in "x is a Brighton bomber" and "x is arrested" but, rather, in some structure (model) in which those things that have been identified as the B things are included in those things that have been picked out as the A things. Hodges' argument implies that only models provide the context and that structures and contexts are essentially the same. It would appear that the necessary conditions for a term to refer involve nothing more than the stipulation of a universe in which the objects are clearly identified.

Hodges' distinction understates the role that the formal language plays. As mentioned earlier, it is not merely the model that allows us to determine how to
evaluate some sentence. It is because our formal language is a first-order language in which the symbols can only play a specific role that our domain makes any sense. Saying that the universe is the set of integers implies that there is a language in which just these objects can be compared with each other in the appropriate manner. The point is not that without language nothing can be compared but that comparisons presuppose commensurable concepts. Without this assumption our comparison of two structures by some common sentence would not be justified. The assumption is that there is a language that captures just what is the same and different in the appropriate manner. (To appreciate that this assumption is problematic, one need only recall Benacerraf’s 1965 article "What Numbers Could Not Be"). Within the minimal meaning/structure distinction a tension exists which threatens to undermine the distinction. The term "minimal meaning" is meant to reflect the contribution that the formal language makes. This contribution is supposed to be of a different kind than that of the structure which is equated with the context of evaluating the referents. If the contribution is semantic, even though it may be of only minimal importance, then the role of the structure is threatened to be diminished. The more content the formal language acquires, the less the model possesses. Furthermore, if the formal language already incorporates part of the semantic component, what prevents one from co-opting more of the role of the context into the language? Thus one might ask why it is

\footnote{In Chapter Two I argued that Frege's criticism of Hilbert's program can be seen to be directed to this assumption and how one captures these similarities and differences.}

75
necessary that the models incorporate the context completely, since our first-order language, with its assumptions, plays an important role in characterizing the context of evaluation.⁶

In the last section of his paper, Hodges introduces another example from English in order to explain the function that non-logical constants play in model theory. Hodges takes the English term "yesterday" to be analogous to a non-logical constant. Contrary to some accounts Hodges argues that "yesterday" does have a meaning. The meaning of "yesterday" is given by a set of rules. The rule he mentions is "the day before today". While Hodges never mentions that the rule itself contains an indexical, he is aware of the problem. Any attempt to give a rule for "today" would necessary involve a referential statement or a new indexical. Instead, Hodges avoids the problem by building in a condition for the use of the rule.

In order to give the word 'yesterday' a reference, we don't have to make a stipulation about how we are going to use it. Rather, we just have to use the word on a particular day, and then the rules of the language decree that it refers to the previous day. (Hodges, 1985/86, pp. 148f.)

It is unclear why the second part of the above quote is more than a stipulation, given the problem that the word "today" raises. What is clear is that Hodges needs this condition to evade Frege's criticism that indeterminate thoughts are neither true nor false. Since indexicals have meanings without having referents

---

⁶ To see why a second-order language is unhelpful here see Chapter Two, specifically my criticism of Boos' argument.
they are indeterminate with respect to reference and not with respect to meaning. Frege’s criticism would then only effect how the rules apply to the elements in a structure. Frege’s objection would miss its mark since meanings are neither true nor false but can still be judged determinate or indeterminate. While Hodges never mentions any other "rules" it is clear that we must assume that what is spoken is English and is not, for example, "Twin Earth English". Hodges’ intention is clear enough: given some rule, together with certain conditions on the use of that rule, we can evaluate the proposition in which the indexical occurs in virtue of the fact that the indexical has a fixed referent in the specific context of use. Given a different context of evaluation the indexical may have a different referent and, thus, may be evaluated differently.

The conditions play an important role in Hodges’ explanation; they obviate the need to give a meaning to the term "today". Clearly the meaning of "today" is not part of the minimal meaning of the language and neither is it included in the meaning that the model or structure provides. While structures provide the determinate elements over which the relations will be assessed, they do not in and of themselves characterize the set of all models which are acceptable. That is, they at best characterize the context in which a reference is possible or not, relative to some intended meaning. Imagine "Twin Earth English" where "today" refers to 100 years of our time. (Their earth rotates their sun once in every 100 years of our time.) It would appear that Hodges’ conditions exclude this context.
In contrast to Hodges one might argue that today contains 100 of our years. This would entail changing what can be said to have occurred yesterday. (That is, it is not true that "Yesterday I walked along the beach".) Disallowing this choice artificially restricts the set of models that could satisfy the theory by only allowing certain kinds of mappings between the theory and the structures. By implication, context and structure are conflated. Thus the rule for applying an indexical, if it is to contribute a semantical component, is part of the structure and "yesterday"’s meaning is completely model-dependent. But then any attempt to explain why only certain mappings between the theory and structure are entertained seems unduly arbitrary. On the other hand, one might stipulate that "today" refers to a 24 hour period of our time from some specific starting point. In the latter case the conditions have dramatically cut up the context of evaluation to suit the language in which the theory has been couched. Hodges takes it to be self-evident that models reflect the possible semantic configurations of the world. The formal theory involves at best a commitment to the minimal meanings of the language.

In his 1922 paper, Thoralf Skolem attempted to show that axiomatic set theory cannot be used foundationally, that it cannot be given a privileged position. Skolem’s account is concerned with the domains over which our theory is to be evaluated. The argument shows that the cardinality of the domain is relative to whether one is looking at the domain from inside the theory or from outside of the theory. What is important to Skolem is not in and of itself that the cardinality of a domain changes with perspective. Rather, set theory as a foundational project
must suppose that the concept of a domain is well understood. Domains are understood in terms of their size and in terms of the relations that hold within them. Skolem's argument is that the cardinality of the domain (when it is sufficiently large) must be assumed prior to the analysis, but then the expectation that set theory will give us an independent account of the concept "cardinality of a set", and thus of number, is problematic. While Skolem's argument is in terms of domains, a related point can be made in my argument. Hodges takes models as canvassing all the semantically significant configurations of the world. This in turn reflects a commitment to the expressive power of the language used. Hodges' argument depends on the relation between some specific language and the set of all significant configurations of the world. Models were supposed to demonstrate their universality in virtue of their freedom from the theory. Since the theory did not contribute any semantical component, the models were free from any restriction invoked by the theory. As the quotation from Tarski demonstrates, the reference to some specific language is crucial. The problem is that one has incorporated a semantical component into the language.

The latter point can be seen to be at work in the criticism of formal proof-procedures by John Corcoran. Corcoran distinguishes between two kinds of arguments which he calls "premise-conclusion" and "demonstrable". In brief, premise-conclusion arguments are ordered pairs \(<P, c>\), where \(P\) is a set of sentences and \(c\) is a single sentence. Regarding premise-conclusion arguments one
is concerned about validity and invalidity; i.e. whether \( c \) is a logical consequence of \( P \).

*If all sentences in \( P \) were true would \( c \) necessarily also be true?*  
(Corcoran, 1972, p. 26)

If premise-conclusion arguments can be characterized by an ordered pair, demonstrable arguments can be characterized by an ordered triple \(<P,R,c>\), where \( R \) is a discourse. Demonstrative arguments are sound when the discourse represents correct reasoning. (Corcoran, 1972, p. 34) Both kinds of arguments are commonplace in philosophy and mathematics and are copiously illustrated in his article. Corcoran argues that these two kinds of arguments involve different kinds of problems. Premise-conclusion arguments are "correct" if the conclusion follows from the premise while demonstrative arguments are "correct" if the conclusion follows by "correct reasoning" from the premises. Validity is often defined formally in the following way:

\[
< P, c > \text{ is valid if there is no argument } < P, c > ^* \text{ having true premises and false conclusion where } < P, c > ^* \text{ is the result of assigning any non empty domain and any appropriate (extensional) meanings to the content words of } < P, c > . (\text{Corcoran, 1972, p. 43})
\]

As Corcoran points out, it is tempting to take this condition, not simply as a necessary condition for validity, but as a necessary and sufficient condition as well. Corcoran argues that demonstrating that the premisses are true and that the conclusion false is dependent on the availability of interpretations (models).

In any case according to the Tarskian definition of validity, the invalidity of an argument depends on the existence of a suitable domain and there might not be "enough" domains to provide "counter interpretations" for all invalid arguments. (Corcoran, 1972, p. 43)
Corcoran's doubts about the Tarskian definition of a valid argument focus on its sufficiency. Clearly if the Tarskian definition of a valid argument was in fact insufficient then, by Corcoran's formulation, there is a "domain" such that it invalidates the argument but is not a member of the set of domains that are canvassed in normal practice. The claim seems to amount to something like, "X is a model and X is not a member of the set of all models". In this form the problem resembles Skolem's argument. In both cases the difficulty arises because the characterization of a concept (domain/model) is taken to be absolute when in fact the characterization is arguably dependent on some specific perspective.

Earlier I had argued that the language, that is our first-order language, already restricts and orders the range of possible re-interpretations. Corcoran makes a similar point early in the paper. In fact Corcoran's argument can be expressed in my terms. The non-logical terms of the theory already begin to determine the range of possible arguments by demanding that any situation involved in testing an argument for validity conforms to a specific kind of situation, namely a model of our first-order language. Clearly this restriction is not a premise-conclusion argument and, thus, not an argument whose basis is formal in the sense expressed above. My primary concern here is not to show that these assumptions are false, but that the assumptions involve a kind of characterization of arguments which already restricts the range of applicability. Thus the assumption that this restricting is performed by only our models is undercut.

\[\text{\footnotesize \cite{corcoran1972} p. 30. There he identifies it as a category mistake.}\]
Hodges separates what I have called the conditions from the context of evaluation. He thus delimits the set of contexts in which the indexicals can be evaluated. When giving the referent of a term, something has to be understood; in the case of the word "yesterday", how to apply the word "today". The choice comes down to where to place the conditions. Hodges is unclear on where they belong. His view seems to regard them as pre-conditions of a formalism and thus an integral part of the formalism. While there is nothing wrong with this move it does undermine the distinction he is attempting to motivate. Recall that a structure is defined by an ordered pair, a domain and a set of relations. Universal quantification is restricted by the domain, otherwise we would not know both how and what the relations were relating. This restriction is to be part of the context of evaluation, the structure, and not part of the "meaning". To this end meaning and reference are separated. It is because of the freedom we gain by not attaching a referent to the non-logical constants, thus leaving open the question of the domains in which the relations are to be tested, that Hodges can claim that our theory allows us, as in group theory, to

classify all groups as commutative or non-commutative regardless of whether anybody has heard of them. (Hodges, 1985/86, p. 148)

Hodges avails himself of the concepts "groups", "commutative" and "non-commutative" in the same way that he avails himself of "today". They become conditions on the context of evaluation and thus part of the restrictions that the formal part contributes to the range of structures that are allowed to satisfy our theory. That is, they in fact restrict the kinds of mappings between the formal
theory and the models that are entertained. Clearly, if these conditions are part of
the meaning of our first-order languages, then a significant semantic component
must be recognized to be involved in the minimal meaning of the language.
CHAPTER FOUR

The primary concern of this thesis up to this point has been to show that content plays a diminishing role in Detlefsen’s defense of Hilbert’s program, and that Frege’s objections focus on this problem. In this chapter I will examine Detlefsen’s argument that Hilbert’s program is not effected by Gödel’s second incompleteness theorem. It will be argued that Detlefsen’s account of the relationship between real and ideal statements effects his defense of Hilbert’s program against Gödel’s second incompleteness theorem. Gödel’s second incompleteness theorem claims of some specific formula which expresses the consistency of a theory, T, that it cannot be proved in T. Following Detlefsen, I will refer to the specific syntactic formula not provable in theory T as Con₇(T). The formula Con₇(T) is said in some sense to express T’s consistency. Detlefsen notes the fact that Gödel’s incompleteness theorem makes a claim about the unprovability-in-T of the specific formula Con₇(T). Since Con₇(T) is said to express a metalevel proposition, namely that T is consistent, Detlefsen argues, in a novel move, that the unprovability-in-T of Con₇(T) does not entail that the metalevel proposition expressing T’s consistency is unprovable in T. He suggests that,

there might still be some formula other than Con₇(T), expressing the same proposition that Con₇(T) expresses, that is provable in T. (Detlefsen, 1986, p. 81)

I will call this postulated formula which expresses the same metalevel proposition as Con₇(T), but which is assumed to be provable in T, Con₈(T).
In what follows two important points will be argued for. The first is that Detlefsen does not distinguish between the metalevel formula, which I will refer to as Con(T), and the proposition expressed by Con(T). In Chapter Two I argued that Frege’s objections can be seen as being directed at the illegitimate importing of content into the metalanguage. Detlefsen avoids this objection by conflating the metalevel formula Con(T) with its propositional content. It is this very move which is denied us at the object language level and which engenders his criticism of the standard argument against Hilbert’s program.

The second point concerns Detlefsen’s so-called stability problem, the problem of showing that,

every set of properties sufficient to make a formula of T a fit expression of T’s consistency is also sufficient to make that formula unprovable in T (if T is consistent). (Detlefsen, 1986, p. 81)

Detlefsen’s explanation of the stability problem rests upon the assumption, that for every provable formula in T, there is an essential set of properties such that, if some provable formula has this set of properties, then we can claim that the cause of the object language formula’s provability is the metalevel proposition. The difficulty with this claim is that, at best, all we can show is that there is a mapping, \( M_2 \), from the formula, \( \text{Con}_b(T) \), to the formula, \( \text{Con}_c(T) \). Since \( M_2 \) does not guarantee semantic content, we cannot be sure that it is the metalevel proposition which has then been proved.

Detlefsen’s re-analysis of the reliability of finitary arithmetic is motivated by Frege’s concern that ideal mathematical statements, since they are meaningless,
cannot function in an inference. Ideal mathematics must therefore be disjoint from real mathematics. Detlefsen is at least partially correct to characterize Frege’s criticism as a loss of reliability. The further analysis of this loss in terms of "epistemic potency" allows for a re-analysis of finitary arithmetic such that the essential kind of reasoning in contentual mathematics is preserved in ideal mathematics. For Frege the question is not one of defining "epistemic potency" divorced from the content on which our reasoning takes place. For Detlefsen the objects on which reasoning takes place are irrelevant. Finitary reasoning involves "primitive recursive arithmetic", and primitive recursive arithmetic is committed to inferences of a minimal kind. Numerals, as such, are understood as types (in contrast to tokens). Detlefsen underplays the intuitive content of the numerals to maximize his criterion of the unity of contentual mathematics and ideal mathematics.

Detlefsen argues that a proof-schema provides the needed motivation for redefining the essence of contentual mathematical reasoning. Since a proof-schema clearly does not exceed finitary reasoning, it must be accepted as a finitary operation, and since a proof-schema already involves a minimal kind of "inference", viz. reasoning on "types", the essence of contentual mathematics is not to be characterized by the manipulation of concrete tokens (computations). (Detlefsen, 1986, p. 54) Having shown that contentual mathematics can be characterized by these criteria, Detlefsen argues that ideal mathematics does not exceed this characterization.
The need for including "proof-schemas" as instances of finitary reasoning can be seen by examining the negation of universal quantifiers over unrestricted domains and the finitary restrictions on this kind of reasoning. Hilbert argues that the "infinite is born" in those existential statements whose content cannot be expressed by a finite disjunction, and in the negation of a general statement, where the variable refers to an arbitrary numeric symbol. (Hilbert, 1926, p. 194)

Hilbert argues that

\( (x)(x + 1 = 1 + x) \)

is incapable of negation from a finitary perspective. (Hilbert, 1926, p. 194)

Thus the Hilbertian finitist divides mathematics into two different kinds of formulas, those

formulas which meaningfully communicate finitary statements [and, the] other formulas which signify nothing and which are the ideal structures of our theory. (Hilbert, 1926, p. 196)

The variable "x" in (1) is here understood as ranging over everything. If we take the logically equivalent statement of the negation of (1),

\( (\exists x)(x + 1 = 1 + x), \)

then the attempt to prove (2) as true involves proving

\( -(1 + 1 = 1 + 1) \text{ or } -(2 + 1 = 1 + 2) \text{ or } \ldots . \)

But (3) expresses (2) if and only if (3) is denumerably infinite. Proving (2) will then involve examining a denumerably infinite number of instances.

---

\(^1\) Instead of "(x)" as the symbol for universal quantification, Hilbert uses an upside down "A".
Hilbert provides a possible solution. It seems possible that we could rely on the principle of excluded middle to show that

\[ (4) \text{ for some numerical symbol } a, \ a + 1 = 1 + a \text{ is false.} \]

That is, if (1) "holds for every symbol, or is disproved for a counter example", then (2)’s truth or falsity can be determined. As Hilbert points out, proving that (4) holds or not depends upon the meaningfulness of (2). But this is the very thing the finitist is attempting to render as meaningless. (Hilbert, 1926, p. 194)

Hilbert’s solution is that (2) is to be interpreted as a hypothetical judgement which asserts something for the case when a numerical symbol is given. (Hilbert, 1926, p. 194)

The existential quantifier "there exists" serves as an abbreviation for

\[ (5) \neg (x + 1 = 1 + x) \text{ for some } x \text{ such that } 1 < x < 1,000. \]

The existential quantifier thus implicitly defines the range of the variable, i.e. it ranges over the finitarily constructed integers up to some constant. Its meaning is grounded in the sense that it is just an abbreviation for

\[ (6) \neg (1 + 1 = 1 + 1, \ 2 + 1 = 1 + 2, \ldots, \ 1,000 + 1 = 1 + 1,000). \]

The letter "x" is here taken to be a schematic variable which in and of itself is meaningless. It is in its function as a representation of the meaningful (6), or some statement like (6) for some finite positive integer in place of "1,000", that (2) is meaningful.

Sentence (1) then represents a convenient shorthand for a denumerable number of axioms to be adopted by our theory as finitarily provable. The intuitions that one wishes to capture are two. First, if (x)Fx is to range over a denumerably
infinite number of objects, then clearly, for any n such that n < \omega, Fn can be proved by some computational procedure. Since, for every n, Fn is finitarily provable, it is not the case that there is an n such that \neg Fn. Second, the "n"s being considered are just iterations of some primitive notation. For example "n" is just "111...[n times]". The apparent transparency of these iterations is supposed to be intuitively clear.

In the case of the second intuition, it is important to realize that the transparency of the operation of iteration is directly related to the content of the objects we are here iterating. The strokes are to be understood as numerals. Because of the lack of complicating content, numerals provide a convenient example of iteration. To say that they lack a complicated content is not to say that no content is present. Rather, it assumes that the criteria for identifying a unit of iteration is clearly given by the context and the notation. Hilbert's ideal mathematics must either down play the content in contentual mathematics, or redescribe it in such a way that the symbols of ideal mathematics are shown to participate within this intuition.

Detlefsen's need to incorporate a "proof-schema" as an acceptable finitary proof involves at least the following two reasons. Frege, in his attack on Hilbert's program, is very much concerned with the value of ideal mathematical reasoning. As Detlefsen points out, Frege's concern is that if there is a specific kind of reasoning germane to mathematics, then the jump from contentual mathematics to ideal mathematics must preserve this kind of reasoning. If this jump does not
preserve this reasoning, then one must explain why the relationship between contentual mathematics and ideal mathematics is reliable when, for example, "bowling" is not. Clearly a full-blown instrumentalist account need not provide a unified account at all. The mere fact that our mathematics (contentual mathematics or ideal mathematics) is a useful tool, for example in physics, is sufficient to explain its reliability. Detlefsen's characterization of Hilbert's ideal mathematics as instrumentalism does not entail Field's full-blown instrumentalist view of mathematics, in which mathematical objects are non-existent and mathematical reasoning is explained in terms of its utility, much like a story is useful in guiding a child's moral development. Any specific story of itself is not essential to the moral development of the child. Even consistency is only a useful criterion and may in fact change with respect to the goal in a story. The instrumentalist is faced with the problem of why one "story" is used and another is not. As Manders points out, the mathematician is constantly making decisions about which "story" to use when a number of "stories" would seem to be adequate. (Manders, 1987) The choice of one "story" over another is in fact taken to demonstrate the internal demands of the field. Reliability is here seen as a necessary but non-sufficient condition for explaining its adequacy.

In Chapter 3 of *Hilbert's Program*, Detlefsen examines Gödel's second incompleteness theorem with respect to Hilbert's Program. Detlefsen finds a

---

2 From now on when referring to "Gödel's incompleteness theorem" I will always implicitly mean Gödel's second incompleteness theorem.
weakness in the usual application of this theorem in arguments which criticize Hilbert’s program. This weakness is exploited by Detlefsen in order to show that, without further assumptions, Hilbert’s program remains uneffected.

Detlefsen’s weakness is located in the inference from the third to the fourth conjecture in his characterization of the standard argument based upon Gödel’s incompleteness theorem. Detlefsen argues that from,

\[
\text{every finitary truth (in particular every truth of the finitary metamathematics of T) can be "expressed" as a theorem of T}
\]

it is incorrect to conclude that

\[
\text{if } \text{Con}_G(T) \text{ is not provable in T, then } \text{Con}_G(T) \text{ does not express a theorem of the finitary metamathematics of T. (Detlefsen, 1986, p. 78)}
\]

Detlefsen’s argument relies upon the following two facts:

(1) that the proof theoretic operations of a specific theory may be inadequate in some way, and

(2) that Gödel numbering (arithmetization) may not be able to capture the content of some other set of sentences which express consistency, but which can be proven in T. (Detlefsen, 1986, pp. 94ff)

Detlefsen’s argument exploits the fact that there may be some way in which one can express the proposition,

\[
\text{T is consistent,}
\]

say using \( \text{Con}_b(T) \), such that \( \text{Con}_b(T) \) can be proved in T even though \( \text{Con}_b(T) \) cannot. (Detlefsen, 1986, pp. 81f) The assumption is that \( \text{Con}_b(T) \) expresses T’s consistency in a way that \( \text{Con}_b(T) \) does not, and that the resources of the
language of our theory, $T$, are adequate for the proof of $\text{Con}_\alpha(T)$ but not for the proof of $\text{Con}_\beta(T)$. Detlefsen argues that $\text{Con}_\alpha(T)$ may include properties which are not essential to every formula expressing the consistency of $T$. Clearly one reason one might believe that $\text{Con}_\alpha(T)$ and $\text{Con}_\beta(T)$ are not both provable in $T$ is by arguing that the language needed to prove $\text{Con}_\alpha(T)$ is inadequate.

Detlefsen draws attention to the semantic relation between the formula $\text{Con}_\alpha(T)$ and the metamathematical proposition expressed by $\text{Con}(T)$. This relationship has two important aspects:

(1) The relationship between the "formula $[\text{Con}_\alpha(T)]$ and the arithmetical interpretation under the standard semantics for $T$." (Detlefsen, 1986, p. 98) (By the "standard semantics for $T"$ Detlefsen just means arithmetic)

(2) The relationship between "the genuine arithmetical proposition which is $[\text{Con}_\alpha(T)]$’s standard interpretation, and the genuine metamathematical proposition which asserts $T$’s consistency." (Detlefsen, 1986, p. 98)

The proof of the metalevel formula expressing $T$’s consistency, $\text{Con}(T)$, is understood to be unencumbered by (1) and (2). Detlefsen seems to imply that, at the metalevel, conditions like (1) and (2) are not present. The supposition is that at the metalevel our language is interpreted, and that the relationship between the syntactic formulation of the proposition expressed by $\text{Con}(T)$ is unproblematic in a way that $\text{Con}_\alpha(T)$ is not. It is unclear what feature the metalevel has that the
language of T does not, but which enables us to characterize a proposition transparently. One feature that arises is that the metalevel allows us to quantify over every formula which expresses T’s consistency. Another condition is that we are now capable of quantifying over the language L itself. Furthermore Con₆(T) and Con₇(T) both express T’s consistency in the language of T, both are specific formulas in the specific language L(T).

Recall Detlefsen’s purpose for introducing Con₇(T). Con₇(T) is said to express the metalevel proposition in a way that Con₆(T) does not. The implication is that given the formula Con(T) in some interpreted metalanguage, Con₇(T) and Con₆(T) both express the same proposition. That is, for any formula in T, which expresses T’s consistency, we can provide the conditions of identity to determine when these different formulas express the same content. That Con₇(T) and Con₆(T) are syntactically different is relevant to whether or not they are provable, but not to what they express. Con₆(T) incorporates specific ingredients which Con₇(T) does not. Irrespective of these syntactic ingredients, our metalevel proposition is at the very least clear enough to allow us to determine how Con₇(T) and Con₆(T) express T’s consistency in T.

Con(T)’s value is given in its independence of the conditions that restrict how a proposition is expressed in T.³ That is, for any language in which T’s consistency is expressed we can, in principle, determine the way in which that

³ To this point we have, following Detlefsen, conflated the meta-level formula, Con(T), with its propositional content.
language expresses the proposition. Thus we can determine that $\text{Con}_o(T)$ expresses $T$'s consistency in a way which prevents $\text{Con}_o(T)$ from being proven in $T$, and $\text{Con}_d(T)$ expresses $T$'s consistency in a way in which the formula $\text{Con}_d(T)$ can be proven in $T$. The metalevel properties needed to characterize the metalevel proposition are thus sufficient to characterize $\text{Con}_o(T)$ relative to $\text{Con}_d(T)$. An immediate problem arises in specifying the strength of our metalanguage. It seems that we must take it on faith that this metalanguage does not itself incorporate any of the dangers exhibited in the object language. Returning to our conditions for the metalanguage we include the following, that it is necessary that the metalanguage allow quantification over languages.

The condition, that the relationship between the syntactic formulation of the proposition expressed by $\text{Con}(T)$ is unproblematic in a way that $\text{Con}_o(T)$ is not, needs further explanation. $\text{Con}_o(T)$ and $\text{Con}_d(T)$ both express $T$'s consistency. Detlefsen argues that our metalevel proposition involves two mappings in order for us to give an interpretation to $\text{Con}_o(T)$: mapping $M_1$, which takes the object language formula to the arithmetical formula; and mapping $M_2$, which takes the arithmetical formula to $\text{Con}(T)$. Together $M_1$ and $M_2$ are to provide a syntactical and finitary procedure for translating any object language formula into the metalevel formula expressing a specific proposition. $M_1$ and $M_2$ are metalevel functions and depend upon the metalanguage and its expressive power. $M_2$ is supposed to guarantee that the semantic content of our metalevel proposition
Con(T) is expressed in Con_{G}(T). The function, M_2, is characterized as the relationship

which Gödel numbering induces between the genuine arithmetic proposition which is [Con_{G}(T)'s] standard interpretation, and the genuine metamathematical proposition which asserts T's consistency. (Detlefsen, 1986, p. 98)

One way of seeing this is to recognize that for a specific formula, K^*, expressing the proposition K in some language L, the syntactical characterization of K in L is given by L(K^*) and L(K^*)'s unprovability does not guarantee K's unprovability. But this just entails that L(K^*)'s syntactic characterization of K is not the whole story. If M_2 were to involve only syntactic considerations it would not be problematic. Yet M_2 must also guarantee that semantic content is preserved, and it is this very condition that is found to be wanting. Hallett explicitly states that Hilbert's intention was to reduce all semantic problems to syntactic problems. (Hallett, 1990, p. 203)

Detlefsen at times refers to the metalanguage as being informal. This is necessary because the metalanguage is assumed to be an interpreted language, i.e. the "meanings" of the terms in this language are assumed to be given. For the time being this will be ignored. None the less, this claim will be understood as implicitly identifying the metalanguage with the semantic content of some natural language via something like "semantic rules". Detlefsen, as the above quotation states, refers to M_2 as mapping a "genuine proposition" to a "genuine metamathematical proposition". If we extend our original object language to include arithmetic, then M_2 becomes a mapping which takes an object language formula to a
metamathematical proposition. Strictly speaking, this is incorrect. The function, \( M_2 \), maps the formula \( \text{Con}_G(T) \) to the formula \( \text{Con}(T) \). The assumption is that \( \text{Con}(T) \) expresses a genuine metamathematical proposition in such a way that the formal properties of our metalanguage do not mimic the problems in our object language. In the first place one should have serious reservations about this claim. As I have already argued in Chapter Two, the move to the metalanguage merely incorporates more syntax. More importantly, the value of the object language is given in the fact that by restricting what can be done in our language, inconsistencies will become more transparent and retractable. Propositional content, e.g. "Truth", is reduced to the manageable property of being True-in-L. Detlefsen assumes that if the syntax of the metalanguage and the metamathematical proposition are conflated, this does not threaten to introduce the paradoxes. Yet this assumption undermines the very concern which led to the introduction of axiomatic systems.

\( M_2 \)'s primary concern is to map a specific object language formula to the metalanguage proposition. If Detlefsen was not concerned with the semantic rules, then reference to \( M_2 \) as semantic relation which takes a formula to a proposition would not be necessary.

---

4 This extension is not really necessary. We could have focused our criticism on \( M_1 \), which performs the same function of mapping a syntactic element to a proposition.

5 See Hallett 1984, pp. 198ff. Hallett here argues that the paradoxes governed the choice of axioms in set theory.
Detlefsen's problem with $M_2$ can be understood in the following manner. Let $\text{Con}_G(T)$ and $\text{Con}_D(T)$ be two different object language formulas, and let $A$ be the proposition "$T$ is consistent", where $A$ is "expressed" by the metalevel formula $\text{Con}(T)$, then

$$M_2(\text{Con}_G(T)) = \text{Con}(T), \text{ and}$$

$$M_2(\text{Con}_D(T)) = \text{Con}(T),$$

where $M_2$ maps a specific syntactic object onto a "propositional-formula". (Detlefsen, 1986, p. 97) Clearly $M_2$ is not sensitive to how the syntactic objects "$\text{Con}_G(T)$" and "$\text{Con}_D(T)$" express $A$. One question that Detlefsen’s program introduces is whether $M_2$ is sensitive to how the two different syntactic elements, $\text{Con}_G(T)$ and $\text{Con}_D(T)$, express $A$. The fact that in the metalanguage $A$ and $\text{Con}(T)$ are conflated means that $M_2$ is not sensitive to how $\text{Con}_G(T)$ and $\text{Con}_D(T)$ differ. By implication, we are left without any explanation of how $M_2$ guarantees that the syntactic elements are being mapped to the genuine metamathematical proposition. Since we know that some syntactic features are ignored, we can not be sure that the desired ones are preserved. Detlefsen argues that $M_2(\text{Con}_G(T))$ expresses the proposition $A$ in such a way that the unprovability of $\text{Con}_G(T)$ may be sheer coincidence. The implication is that $\text{Con}_G(T)$’s unprovability in $T$ may not be a function of what it expresses. $M_2$ maps $\text{Con}_G(T)$ to $A$ but the syntactic elements of $T$ prevent $\text{Con}_G(T)$ from being proved.

In what follows, the latter claim will be explored. Detlefsen argues that

$X^*$: This problem [that $\text{Con}_G(T)$ is not provable in $T$] is prompted by the possibility that the properties of $\text{Con}_G(T)$ which Gödel’s proof calls
upon to show the unprovability-in-T of Conₐ(T) may not all be included among those properties of Conₐ(T) which cause us to say that it expresses the consistency of T. (Detlefsen, 1986, p. 81)

Generalizing X* we get

X: For any specific object language formula of T, the proof of that formula in T may depend upon its formal properties and not upon the properties which cause us to say that it expresses the metalevel proposition.

Recall the problem with M₂, i.e. with the relationship between the standard interpretation of Conₐ(T) (the arithmetical formulation) and the proposition "T is consistent". Detlefsen argues that

Y*: if the Gödelian is to find a solution to the Stability Problem for a given system T ... he must locate a set C of conditions on formulae of T ... such that

(1) every formulae of T that can be reasonably be said to express the consistency of T satisfies the conditions in C, and

(2) no formula of T that satisfies C can be proven in T provided T is consistent. (Detlefsen, 1986, p. 93)

Recall that the stability problem is to show that every set of properties sufficient to make a formula of T a fit expression of T’s consistency is also sufficient to make that formula unprovable in T (if T is consistent). (Detlefsen, 1986, p. 81)

It is important to notice that Y* is an essentialist claim. It identifies a set of sentences by a set of properties essential for their proof in T. Detlefsen’s Y* makes a general claim about "unprovability". Generalizing Y* and the "stability problem" to the case of provability, he is left with the following claims: The stability* problem is to show that
every set of properties sufficient to make a formula of \( T \) a fit expression of some metalevel proposition of \( T \) is also sufficient to make that formula provable in \( T \) (if \( T \) has that property).

\( Y^* \) then becomes,

\[ Y: \text{In order to find a solution to the stability}^* \text{ problem for a given system } T, \text{ Detlefsen must locate a set } D \text{ of conditions on formulas of } T, \text{ such that} \]

\( (1) \) every formula of \( T \) that can be reasonably be said to express some metalevel proposition \( D \) satisfies the conditions in \( D \), and

\( (2) \) no formula of \( T \) that satisfies \( D \) is unprovable in \( T \), provided \( T \) is consistent.

As it stands, \( Y \) does not clearly express Detlefsen's concern expressed in \( X \). That is, conditions (1) and (2) in \( Y \) do not guarantee that the conditions given in \( D \) are those properties which allow us to say that the proof of an object language formula is due to the proposition it is expressing and not to some accidental formal properties of our object language. Thus to \( Y \) we add

\( (3) \) the set \( D \) of properties are those properties which allow us to say that for those formulas of \( T \) which express \( D \) and are provable in \( T \), are provable because of \( D \) and not because of some accidental feature of \( T \).
What ever the set $D$ is, it must be constitutive of finitary arithmetic. Two important problems arise. Clearly, the best explanation for $D$ is the set of properties which demarcate finitary arithmetic. The danger of begging the question is immediately apparent. The second problem is specifically due to Detlefsen’s characterization of ideal mathematics as meaningless. If the set $D$ of properties is to be of any use to him, it must demarcate the set of finitry arithmetic as a completed set. That is, the set $D$ must be impredicatively defined, and thus $D$ must exceed the very restrictions Hilbert places upon what the universal quantifier can finitarily mean.
CONCLUSION

Michael Detlefsen provides us with a Hilbertian account of mathematics. This account respects Hilbert's finitary/ideal distinction. Detlefsen's characterization of the evaluation of ideal proofs forces him to treat the content of both finitary and ideal mathematical statements as consisting in forms. I have argued that this move should result in our questioning the intuitions involved in finitary reasoning. Hilbert regarded these intuitions of great importance when assessing the reliability of a proof. Furthermore, I have argued that Detlefsen's account threatens to relativize the concept of "being finite". Frege's critical remarks show the need to regard the conceptual framework in which a mathematical solution is sought as crucial to understanding his objections to Hilbert's independence proofs. While Detlefsen is aware of Frege's intent, he avoids the problem by conflating the metalevel formula with the metalevel proposition it expresses. This conflation need not be seen as a problem in general. Since Detlefsen's criticism of the standard argument, in defense of Hilbert's program, relies upon the separation of a syntactic formula and the metamathematical proposition it is said to express, he leaves unmotivated the conflation at the metalevel.

Frege's account is understood as questioning the notion that an uninterpreted theory is capable of performing its expressed purpose. In Chapter Three, Frege's objections are applied to Hodges' defense of Hilbert's model.
theoretic program. I argued that the background language of set theory is presupposed in order for an indexical to function as it is assumed to function in a natural language. In natural languages, content is assumed. Frege's critical remarks to Hilbert question the assumption that an account of a mathematical theory can be given without also giving an account of the background language, i.e. the conceptual framework, in which mathematical problems live.

While Detlefsen's excellent book *Hilbert's Program* provides us with an interesting development of Hilbert's position, Detlefsen fails to develop adequately the relationship between finitary mathematics, ideal mathematics and the metalanguage. One aspect of this failure is that Detlefsen's account of why some formula may not be provable, does not in any apparent way suggest how one should construe the concept "proof-in-T". I argued that this failure is in fact a function of the separation, in the object language, of the formula and its propositional content. For all the failures of Frege's formal theory, he was aware of this problem and critical of Hilbert's program because of this separation.
BIBLIOGRAPHY


-------- [1879] "Begriffschrift, a formula language, modeled upon that of arithmetic, for pure thought". Reprinted in van Heijenoort [1967], 1-82.


106


