1994

Term-forming Operators In First Order Logic

David Michael DeVidi

TERM FORMING OPERATORS IN FIRST ORDER LOGIC

by

David DeVidi

Department of Philosophy

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
March 1994

ISBN 0-315-93195-7

ABSTRACT

The two main accomplishments of this thesis are that it provides the first adequate semantics for Hilbert's epsilon–operator and that it describes a general semantics for term forming operators (often called "variable binding term operators" or "vbto's") more flexible than any in the literature.

The epsilon–operator was introduced by David Hilbert in the 1920s as a term forming operator in first order logic. The semantics so far available for epsilon has been designed for classical two-valued logic, and has required that additional extensionality assumptions be made. This thesis provides complete semantics for epsilon in classical extensional, classical non-extensional, Boolean valued, and intuitionistic first order systems. The natural step to generalizing the techniques used in the epsilon case to get a general theory of term forming operators which handles the non-extensional and non-classical cases is then taken.

The thesis proceeds as follows. Chapter One gives a historical discussion of term forming operators. A brief, self-contained presentation of the untyped lambda-calculus, which illustrates the inevitable differences between lambda and any possible operator in first order logic, follows. A chapter is devoted to solving the syntactical difficulties involved in introducing a variable binding term forming operator to standard languages for first order logic. The semantics for epsilon, and in the intuitionistic case also for another of Hilbert's creatures, tau, takes up the next several chapters. The discussion includes several new completeness and soundness results, and some new results about the extra strength these operators add to intuitionistic logic, including some new independence results. The final chapter includes an argument to the effect that the results earlier in the thesis show that we need a more general general theory of term forming operators than any in the literature, and indicates the shape such a theory should take.

FOR JANE,[1] WITH LOVE

[1] Though what good it is to her is beyond me!

ACKNOWLEDGEMENTS

First of all, to keep the police of my back and for other better reasons, I should mention that this thesis was typeset using $\mathcal{AMS}$-TeX.

In my overly long career at Western, which finally ends with this thesis, I have incurred many debts, some professional, some personal, and many financial. I would like to take this opportunity to acknowledge, however inadequately, some of each of the first two sorts.

First I want to thank two people in particular for their work as teachers. John Bell has been a veritable fountain of ideas, information, inspiration and advice, not the mention his having demonstrated remarkable patience in serving as thesis supervisor to someone like me. Bill Demopoulos runs amazing courses: ones for which it is worthwhile to go back and read the lecture notes years later. I was lucky to be able to take several of these. Most of all I want to thank these two because, each in his own way, they have supplied me with encouragement and between them they convinced me, rightly or wrongly, that I had the talent to work on the things that really interested me.

On a more personal level. I would like to thank all my office-mates over the last several years, especially Rob, Len, Ian, Philip and Rick (an honourary office-mate for all the time he has spent there) for putting up with me, for being really smart guys worth talking to about just about anything, and generally for being such good company. John and Mimi Bell supplied me, on many an occasion, with great food and even better company. Life at Western improved immensely with their arrival in London. I am grateful to all the people who have served on the support staff in the Philosophy Department for making sure that there was always some place in the department where even nobody graduate students were treated like people, and especially to Pauline Campbell for more than once pulling me back from the bureaucratic abyss.

It seems that life in grad school turns everyone into a self-centred jerk. I want to thank my family, especially my mother, my father, Rob and Rick, Grant, Darlene, Wendy and Jennifer for pretending not to notice. Most of all, I want to thank Jane for reasons too numerous and personal to include here.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION: SOME HISTORY, SOME COMMENTARY

An undergraduate career in philosophy cannot last for long before running up against a discussion of the definite description operator ι and its seemingly inseparable travelling companions 'the present King of France' and 'the author of Waverley'. If our student is brave enough to take a course in the Philosophy of Mathematics, she or he is bound to be told that a crisis in the foundations of mathematics came about when the paradoxes, Russell's paradox in particular, showed that the *naive conception of set* on which "any predicate has an extension" is inconsistent. Luckily, though, the world did not have long to wait before Zermelo (among others) came along to invent a set theory which, among all the set theories so far proposed, "alone ... is not only consistent (apparently) but also an independently motivated theory of sets."[1] A crucial feature of this set theory is that we cannot allow unrestricted application of the operation of forming extensions, but instead must adopt the Axiom of Separation. That is, we cannot simply have an unrestricted rule which says that $\{ x \mid \varphi \}$ is a set, but instead can only say that *if y is a set* then $\{ x \in y \mid \varphi \}$ is a set. In this course our student may also run up against Hilbert's ε operator.

[1] These quotes are from [Boolos].

The single axiom for ε is supposed to allow us to derive all the axioms that govern reasoning about the transfinite, and so ε is the most important tool we have for ensuring that "no one shall be able to drive us from the paradise that Cantor has created for us" [Hilbert 1925].

What the description operator, the abstraction operator and ε have in common is that they are all, from the point of view of formal logic, what we shall here call *term forming operators*. We shall mean by this that in a formal language $\mathcal{L}$, for each variable x and each formula φ of $\mathcal{L}$, application of the operator gives us a *term* of the language which in these cases are denoted by '$\iota x.\varphi$', '$\{ x \mid \varphi \}$', and '$\varepsilon x.\varphi$' respectively; furthermore, if x was free in φ, it is no longer free in the term resulting from the application of the operator. The official topic at hand in this thesis is the behaviour of such operators in first order logic. What I hope to do in the present chapter is both to introduce the subject and explain what I have included and what I have excluded.

These operators are, from the point of view of ordinary first order logic, strange hybrids. For in the usual presentations of first order logic, the closest thing to something we might call a term forming operator is a function symbol. We can describe an n-ary function symbol f as an operator which, applied to an n-tuple of terms $\langle t_1, \ldots, t_n \rangle$, returns a term $f(t_1, \ldots, t_n)$. But the operator here is applied to (sequences of) terms and not to formulas, and there is no variable binding involved. Where there is variable binding in standard presentations of first order logic it is done by the quantifiers, and the output of these operators is a formula, not a term.

1. Description Operators

Term forming operators, in the sense we intend here, have been around since the start of the century. Newton da Costa, in his brief history of the general study of term forming operators [da Costa 1983], dates the "first solid hint of awareness" of such operators to [Padoa], which includes a brief discussion of appropriate formalizations of the definite article. For some unstated reason da Costa does not include the "substitute for the definite article" $\backslash\xi$ of [Frege]. The first volume of *Principia Mathematica* contains a well-known discussion of the description operator, as do several other of Russell's writings from the first quarter of the century.

Description operators were not only the first term forming operators to be discussed, but they have since that time been by far the most discussed of term forming operator in the philosophical literature. Frege's name for it makes obvious why this is the case: the description operator is the obvious formal language analogue of the definite article of natural language. Discussions of ι therefore have an obvious relevance to ordinary language reasoning and, conversely, disputes about the truth of claims about the present King of France can have implications for the proper form a formal treatment of ι ought to take. Since we want $\iota x.\varphi$ to refer to the unique object x with the property φ, there is not much to be said about the term when there *is* such a unique object. Predictably, then, most of this discussion has centred on what ought to be done with improper terms –that is, it has centred on the question of what to do if the number of objects with the property φ is other than one.

Russell's approach to improper definite descriptions is (to a first approximation) to make all statements in which such a description occurs false. Formally speaking, his approach was to regard the description operator as not, strictly speaking, part of the object language. Rather Russell's description terms are merely abbreviations of more complicated expressions of the language, and the expressions abbreviated are such that sentences including the descriptions are false if the description is improper. Objections to Russell's contextual definitions for ι usually took the form of pointing out that they gave counter-intuitive results when translated into an ordinary language. For instance, it might be claimed that it is true that the golden mountain is a mountain, or that Sherlock Holmes played the violin (after a fashion).

Those interested in building a formal logic with descriptions are also likely to have general objections to treating ι–terms as abbreviations, not just objections to the specific details of Russell's theory. For it is, formally speaking, quite inconvenient. For one thing, it is rather awkward to have the rules of inference not apply directly to expressions containing these terms, but, if the terms are not really part of the object language, they must be "deabbreviated" before inferences can be made [Corcoran and Herring, p. 648]. Attempting to circumvent this difficulty "leads quickly to elaborate technical maneuvers involving numerous metatheorems in order to justify their use in a natural and easy way [Corcoran, Hatcher and Herring, p. 177]."

(Of course, the contextual definition approach to term forming operators has its defenders. The great virtue of the contextual definition approach is that it ensures the eliminability in principle of these terms, and so circumvents the need for an independent (formal) semantic analysis of them. This is part of the appeal of this approach for Quine, one author who continues to use it, e. g. in [Quine].)

It is presumably considerations of the second sort that have led many authors, going all the way back to Frege, to take the description operator as genuinely a part of the object language,[2] and not as merely an abbreviatory device. This results, of course, in a need to give a semantic treatment of the same sort to these terms as is given to the other components of the language under consideration. Since there is not a lot of room for interesting variation in the case of proper ι-terms, it is in the treatment of the improper ι-terms that such semantic treatments vary. We can divide the varying theories, in a rough and ready way, into two types.

The first, often called a Fregean approach, is to simply have all the improper ι–terms denote a particular null-element in the domain.[3] The great appeal of this is, of course, its simplicity. This, together with the obvious interpretation of proper terms, immediately determines the semantical interpretation for all the description terms. If in addition to our favourite axiom to handle the behaviour of proper ι–terms we also add the axiom $\neg\exists!x.\varphi \Rightarrow \iota x.\varphi = \iota x.x \neq x$, we can easily prove soundness and completeness.

The objections to this approach are not hard to imagine. For suppose we really want ι to be a formal analogue of the definite article. The Fregean approach is obviously determined by considerations of technical expediency, and the result is that, if we adopt the supposed point of view, we must conclude that the formal tail

[2] To speak, I think harmlessly, a bit anachronistically for Frege.

[3] I count as Fregean those approaches, like the one in [Bernays and Fraenkel], which in effect build a null entity right into the language.

is wagging the natural language dog. For why would we want to have 'the present King of France' refer to an existing object, though admittedly an arbitrary one, when what makes it improper is precisely that there is no existing object for that term to refer to? And while the money tree in my back yard and my classic good looks are, unfortunately, both non-existent, it seems strange to conclude on that account that they are the same thing.

Such considerations have led a number of authors since the early 1960s to move towards a semantics where improper terms simply do not refer to any element in the domain of quantification. Indeed, the unnaturalness of the Fregean approach to improper descriptive terms is one of the key motivations (along with the desire to eliminate the need to suppose the domain of interpretation to be non-empty) behind the developments in the growing literature on free logic. There are many different proposals as to what shape a semantics for free logic should take, many of which are defended by their authors as capturing (at least some aspect of) the behaviour of improper descriptive terms in ordinary language. A speedy resolution of these debates seems unlikely. For some useful discussions of description operators in free logic, see [Scott, 1967, 1970], [Bencivenga et al.], and the papers in [Lambert 1970, 1991].

2. Other Operators

While the description operator has been much discussed by philosophers precisely because it is such a close analogue of a familiar feature of ordinary language, it is of, at best, minor mathematical significance. Other term forming operators have by and large had to earn their keep by being of some mathematical use.

Some of these operators have earned their keep by being key features of (some formulations of) certain "special sciences". I have in mind in particular set abstraction, which is at home, obviously, in set theory, and minimization in number theory. Each of these operators has, of course, been much studied precisely because it plays an important role in a heavily studied area of mathematics. We will have little to say about such operators here. Instead, our focus will be on those operators (like ι and ε) which can be regarded as *logical* in the sense that we can regard the operator as part of the basic structure of our language. To put the point another way, we want to deal with operators which can be counted among the logical constants of our languages because we want our languages to be general purpose first order languages, not languages which are appropriate only to one or another area of mathematics. This is not as significant a restriction as it might at first seem, since it is often not terribly difficult, by careful selection of axioms, to make a term forming operator which is taken to be a logical constant behave in a manner suitably close to a special purpose operator like minimization. For some examples, see [Leisenring, ch. 3–4]. More detailed treatment of operators which are specific to a particular science is a proper subject of research for workers within that science.

Term forming operators which we might call *logical* in the sense just indicated have also been in use since early in the century. In 1923 Hilbert introduced his τ operator, whose behaviour is governed by what he called the *transfinite axiom*, which we will simply call the (τ)–axiom,

$$(\tau) \qquad \varphi[x/\tau x.\varphi] \Rightarrow \forall x.\varphi.$$

(Here $\varphi[x/\tau x.\varphi]$ is the formula which results if we replace all free occurrences of x in φ by $\tau x.\varphi$, renaming bound variables as necessary. This will be made precise in Chapter 5.) In [Ackermann], which was a PhD thesis written under Hilbert, ε makes its first public appearance. And λ, which we will discuss below, became a subject of intense research by the mid-1930s. It will be useful to say something about how these operators were supposed to earn their keep.

3. Hilbert's Operators

In 1925 Hilbert pronounced that "the situation in which we presently find ourselves with respect to the paradoxes is in the long run intolerable. Just think: in mathematics, this paragon of reliability and truth, the very notions and inferences, as everyone learns, teaches, and uses them, lead to absurdities. And where else would reliability and truth be found if even mathematical thinking fails?" As a consequence, "the definitive clarification of the *nature of the infinite* has become necessary, not merely for the special interests of the individual sciences, but rather for the *honour of the human understanding itself*" [Hilbert 1925 pp. 375, 370–71]. The "clarification" Hilbert had in mind was in fact to eliminate *the infinite* from the catalogue of things regarded as real, and so to restrict its role to "merely that of an idea." He meant by this that the infinite is to be regarded as a mere device, the use of which allows us to continue to operate with as formally simple a set of rules as possible. In many cases if we were to strictly eschew mention of the infinite in our reasoning we would be stuck with arguments of great complexity and little perspicuity, assuming we could arrive at a proof at all. Appeal to the infinite allows us to discover and present proofs that are clear and straightforward. However, this use of the infinite could, according to Hilbert, only be secured by means of the finite. In short, Hilbert's view was that it was necessary to establish that anything that can be proved about the actual and finite by reference to the infinite can also be proved without such reference being made.

Both τ and its replacement ε were supposed to play a central role in this "definitive clarification". The *logical ε-function* was governed by the following axiom.

$$(\varepsilon) \qquad \varphi \Rightarrow \varphi[x/\varepsilon x.\varphi].$$

The virtue of either of these operators is supposed to be that it allows us to define the classical quantifiers, ε, for example, doing so according to the following schemes:

$$\forall x.\varphi \iff \varphi[x/\varepsilon x.\neg\varphi]$$

$$\exists x.\varphi \iff \varphi[x/\varepsilon x.\varphi].$$

These operators were therefore supposed to allow us to derive all the classical modes of reasoning about the infinite, and so to defend the simple modes of inference due to Aristotle. Hilbert was particularly concerned to defend what he called the law of excluded middle,

$$\neg\forall x.\varphi \Rightarrow \exists x.\neg\varphi,$$

from the carping of intuitionists and other constructivists. (See especially Hilbert's diatribe in [Hilbert 1927, pp. 472–76].) Presumably Hilbert abandoned τ in favour of ε because he regarded ε as more transparently a correct principle. In either case, basing all reasoning about the infinite upon a single principle was intended to make it easier to supply the finitary consistency proofs that would guarantee that our use of the infinite will not lead us into error.

Hilbert's programme in the foundations of mathematics rather famously came a cropper with the publication of Gödel's incompleteness results. More interesting for our present purposes, though, is another problem. By 1925 Kolmogorov had shown that one of Hilbert's propositional axioms, the law of double negation

$$\neg\neg\beta \Rightarrow \beta,$$

when added to the *intuitionistic* first order predicate calculus, already allows us to prove the classical rules Hilbert was so concerned to defend. Moreover, it has recently been shown (in [Bell 1993a]) that if we add ε but not the law of double negation to first order intuitionistic logic we *cannot* prove those principles. So even if Hilbert had succeeded in convincing some intuitionists that the ε–operator embodied valid principles of reasoning, he could not have gone on to prove that the classical reasoning patterns were acceptable without first convincing these intuitionists of the acceptability of the law of double negation. Since that law is equivalent to the (genuine) law of excluded middle

$$\beta \vee \neg\beta,$$

this last possibility is, to say the least, remote.

In spite of these failures, however, ε has proved to be mathematically useful in a variety of ways. [Hilbert and Bernays vol. 2] includes a comprehensive account of the results attained by the Hilbert school using ε, including consistency results for some restricted areas of mathematics. A brief account of some of these results can be found in [Leisenring]. ε has also proved useful as a tool for giving particularly elegant presentations of specific theories. For instance, [Bourbaki] contains a presentation of set theory which takes ε as a basic logical symbol, along with $\in$. So while its role in foundations did not turn out to be what Hilbert wanted it to be, ε has turned out to be mathematically useful in its way nonetheless.

4. Church's λ–Operator

In a small, unscientific sampling of philosophers, namely those with whom I have discussed the subject matter of my thesis, mention of term forming operators most often brought to mind the description operator ι and the abstraction operator λ of Church's λ–calculus. The λ–calculus is the subject of an enormous literature, especially if we include the literature on its (as it turns out) close relative, combinatory logic, so it is perhaps not surprising that more philosophers would have heard of λ than, say, τ.

From the point of view of this thesis, the λ–operator is an interesting case. Part of the interest in the λ–calculus is that it is a vehicle for the study of variable binding and substitution in general, and one of the better known results of the study of λ–calculus is that any variable binding syntactical operator can be defined in terms of λ and a suitable non-variable binding operator (see [Curry and Feys, pp. 85–86] for a proof). So in a sense the λ–operator can be viewed as the only all-purpose variable binder we could ever need.

But the λ-operator is not a term forming operator of the sort announced above as the official subject matter of this thesis. For the λ-operator is at home only in a term calculus—that is, in a calculus in which all the well-formed expressions are terms. To put the point another way, while the output of the λ-operator is a term, just as it is for ι, τ and ε, its input is different. These operators take a variable and a *formula* as input, while λ takes as input a variable and a *term*. Now, it is well-known that with suitable ingenuity a language for type theories can be formulated as a term calculus (for an elegant example, see the presentation of type theoretic languages for local set theories in [Bell, 1988]), so λ can be called a term forming operator in type theory with as much justification as ε, ι, or $\{:\}$. But since our concern here is first order logic, we cannot simply dismiss the distinction between terms and formulas.

The reason for λ's failure to qualify as a term forming operator in first order logic is of some interest. The λ-calculus was originally intended by Church to be an analysis of the notion of "functions in general", and so was to play an important foundational role for mathematics. The λ-calculus, as originally conceived, was supposed to be an *untyped* calculus. Unfortunately, it did not take long for people to show that the untyped λ-calculus, while it was proved to be consistent as early as 1936, was not consistent if we attempted to add either negation or implication. This left workers with a choice between accepting the limitation to a logic-free calculus or moving to a typed calculus of some sort.

Since the syntactical investigations of λ-calculus have been thorough, and since one of the points of such study has been to conduct a general investigation of variable binding, we might expect to learn some general lessons about the proper syntactical handling of term forming operators in logic from handling of the λ-operator. Also, since λ seems to be more familiar than the other operators we will be concerned with, it will be worthwhile to make as clear as possible its relationship to these other operators. So Chapters 3 and 4 below will be devoted to the untyped λ-calculus. However, we will be careful to refer to that calculus as *untyped*, while referring to the calculi we will investigate in later chapters as *first order*, and we should note that because of this a more proper description of the subject matter of this thesis would be to say that it is *term forming operators in untyped calculi*.

5. Formal Semantics for Term Forming Operators

The mathematical uses of term forming operators up to the 1940s took place exclusively within a proof-theoretic context. This is not to say that there were no discussions of the *meaning* of these operators, as is reflected, for example, in the criticisms of Russell's account of the description operator described above. Rather this simply reflects the fact that until the work of Tarski and others opened up the science of formal semantics there was essentially no framework for a formal semantical treatment of term forming operators. But once formal semantics started to become a central part of logic it did not take long before formal semantic treatments of specific operators began to be proposed.

Not surprisingly, proposals for a semantic treatment of ι were the first off the mark. As was mentioned above, the various proposals were distinguished by their treatment of improper terms. For a discussion of various early proposals, see the discussion in [Carnap 1956]. In the early 1960's proposals for a semantics for ι in free logic were developed by a variety of authors.

The first thorough semantic treatment of ε was presented in 1957 in [Asser]. Asser actually presented three different methods for constructing a semantics for ε, two of which we will present below. The third semantics, which Asser suggests is too complicated to really capture what Hilbert must have had in mind, makes the interpretation of an ε–term $\varepsilon x.\varphi$ depend not only on the interpretation of φ, but also on the syntax of $\varepsilon x.\varphi$.[4] His second semantics could be called a first attempt at a free-logic semantics for ε, since an $\varepsilon x.\varphi$ does not get an interpretation on this semantics unless we can prove $\exists x.\varphi$. (So this is more properly regarded as a semantics for another of Hilbert's creatures, the η–operator. We will have more to say about η in Chapter 9.) His first and "preferred semantics" was one which took Hilbert's comments about the relationship between ε and the axiom of choice seriously, and so interpreted the ε–terms using a choice function on the power set of the domain of interpretation in as direct a way as possible.

Asser's idea is a natural one which will continue to play an important role in our discussion, so it is worth describing the essentials of his idea in a bit more detail. We write '$\|\varphi\|^{\varrho}_{\mathcal{M}}$' for the truth-value of the formula φ in a structure $\mathcal{M}$ under ϱ, where ϱ is a map taking the set of variables into the domain D of $\mathcal{M}$. The map $\varrho(x/d)$ agrees with ϱ except possibly at x, where it assigns the value $d \in D$. The *truth-set*

[4] See Chapter 8 §1 for a detailed presentation of a semantics which captures the essentials of this proposal.

for φ and x (for the interpretation determined by $\mathcal{M}$ and ϱ) is the subset of the domain D such that $\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ is *true*. We can then interpret ε-terms by adding a choice function $E : \mathcal{P}(D) \to D$ to our structure, then having the interpretation of $\varepsilon x.\varphi$ (under ϱ and $\mathcal{M}$) be the element chosen by E from the truth set for φ and x (under ϱ in $\mathcal{M}$).

This is very a simple and appealing idea, and it is not hard to see that this approach will ensure that the (ε)–schema is valid. However, it is also not hard to see that this approach will have other consequences. (1) First, and rather unobjectionably, $\varepsilon x.\varphi$ and $\varepsilon y.\varphi[x/y]$ will get the same interpretation. This is probably a result that we want, since it conforms with our usual understanding of appropriate behaviour for bound variables.[5] (2) More significantly, any two predicates with the same truth set will give rise to ε–terms which get the same interpretation. The result is that in order to get a calculus which is complete for this semantics we must add the following scheme to our list of axioms:

$$\text{(Ack)} \qquad \forall x.(\varphi \Leftrightarrow \psi) \Rightarrow \varepsilon x.\varphi = \varepsilon x.\psi.$$

Asser admits that his motivation for accepting this schema is simply that it is required by his straightforward semantics. [Leisenring p. 33], the only book-length treatment of ε, simply adopts Asser's preferred semantics and rather tendentiously specifies that one condition on a "suitable" semantics for ε is that it must make (Ack) valid. We will see many reasons to reject this stipulation below.

[5] Asser has some unusual syntactic restrictions built into his system which eliminates the need to mention this point.

It is not hard to see that a Fregean semantics for ι can easily be described in terms of truth-sets. We can interpret ι by adding to a usual first order structure a function $f : \mathcal{P}(D) \to D$ which is a choice function on singletons, but which maps all non-singletons to a fixed but arbitrary element of D.

This suggests a way we might move towards a general theory of term forming operators. The terms formed by a term forming operator σ are interpreted by functions $f : \mathcal{P}(D) \to D$. If M is the truth-set for x and φ (in $\mathcal{M}$ under ϱ), the interpretation of $\sigma x.\varphi$ (in $\mathcal{M}$ under ϱ) is $f(M)$. We can regard the identity of different operators as being determined by the sort of maps $f : \mathcal{P}(D) \to D$ we allow to interpret the operator. The first to suggest this approach to giving a general theory of term forming operators seems to have been Dana Scott in [Scott 1967], where he points out that a consequence of this simple theory would be the need to adopt axioms which take account of (1) and (2) above.

Unfortunately Scott's work seems to have gone unnoticed by many later researchers. The first sketch of a general theory of term forming operators which seems to have generated much attention seems to have been the one given in [Hatcher 1968].[6] Hatcher interprets our terms by a function f which takes into D (not simply subsets of D but) finite sequences consisting of elements of D and subsets of the domain D. The interpretation of the term formed by our operator from x and φ is defined to be the value of this function on the sequence $\langle r_1, \ldots, r_n \rangle$, where if $\langle x_1, \ldots, x_n \rangle$ is the sequence of free variables in φ, $r_i = \varrho(x_i)$ if x_i is not x,

[6] Hatcher introduces the inelegant name *variable binding term operator* for what we have been calling term forming operators here, and that name seems unfortunately to have stuck. I will nonetheless continue to use the name *term forming operators*.

and r_i is the truth set for φ and x if x_i is x.

[Corcoran and Herring] raised a number of objections to Hatcher's proposal. They object that Hatcher's theory has the consequence that the rules of Universal Instantiation, Existential Generalization and Substitutivity of Identicals no longer hold in full generality. They also object that the interpretation of a term depends on the order in which, for instance, variables occur in that term even in cases where this seems inconvenient—for instance, the terms formed by binding x in $x = y$ and in $y = x$ could turn out to be different. Their proposal is to solve these problems by making the interpretation of a term depend only on the truth-set, and they make the conjecture that if we then adopt a principle they call the truth-set principle, we will be able to prove that the resulting calculus is complete for their proposed semantics. The truth-set principle is the scheme that, if x but not y is free in φ and y but not x is free in ψ, then

$$\text{(TSP)} \qquad \forall x.\forall y.(x = y \Rightarrow (\varphi \Leftrightarrow \psi)) \Rightarrow \sigma x.\varphi = \sigma y.\psi.$$

Shortly thereafter Corcoran and Herring teamed with Hatcher to prove that their conjecture is correct in [Corcoran, Hatcher and Herring]. (Their proof is sketched in Chapter 12 below).

The complaints raised by Corcoran and Herring seem to have been widely accepted. It certainly seems to have convinced Hatcher, who in his later book [Hatcher 1982] presented a general theory of term forming operators very much in the spirit of Corcoran and Herring. Newton da Costa pointed out almost immediately in his review [da Costa 1973] of [Corcoran, Hatcher and Herring] that (TSP) is equivalent to the conjunction of (Ack) and a rule allowing us to rename variables bound by a term forming operator. So, in effect, Corcoran et al. had rediscovered the semantics suggested by Scott, which was the obvious generalization of the preferred semantics for ε presented by Asser in 1957, neither of which they seem to have been aware of. This general theory has been the subject of a lot of investigation by da Costa and his collaborators since; see [da Costa 1973, 1980], [da Costa and Mortensen], [Druck and da Costa]. These researchers have succeeded in showing that much of standard first order model theory carries over to this new semantics.

All of the complaints raised by Corcoran and Herring arise because Hatcher's original semantics does not guarantee that we will be able to substitute equivalent formulas or identical terms within the scope of a variable which is bound by a term forming operator. In effect, then, their complaint is that Hatcher's semantics creates terms which do not behave in an extensional enough manner. But it is by no means obvious that we ought to want our term forming operators to behave in a strictly extensional way. For example, if we simply look at the (ε) scheme, it is natural to regard $\varepsilon x.\varphi$ as not just a φ, but as the *paradigm* φ, or the *ideal* φ. (The first reading is attributed by [da Costa and Mortensen] to Malcolm Rennie. The second is used in [Bell 1993a].) On this reading there is no reason to think that the fact that two predicates are co-extensive should mean that they give rise to the same ε–representative. To borrow Bell's example, there is little reason to expect that the ideal human and the ideal featherless biped will be the same. But there are more compelling reasons for not simply accepting the extensional theory proposed by Corcoran et al. as a general theory of term forming operators, and we will return to them below.

6. Generalizing the General Theory

The discussion of the theory of term forming operators so far has been almost exclusively a discussion of term forming operators in classical two-valued logic. But as we saw Hilbert's goal was to use the ε–operator to justify the use of classical patterns of reasoning in the face of objections from intuitionists. It would seem natural, then, to try to investigate the effect of adding ε to *intuitionistic* predicate calculus. There are other factors which make an investigation of term forming operators in intuitionistic logic a pressing matter, notably the fact that the development of Topos Theory has moved intuitionistic logic from the fringes of logic into the middle of the action.

Hilbert's *Second ε–Theorem* in effect tells us that the addition of ε to *classical* logic is conservative in the sense that anything that can be proved using ε, starting from an ε–free set of premises and finishing with an ε–free conclusion, can also be proved without it. While it has long been known that ε is non-conservative in the intuitionistic case (for example, it obviously makes the intuitionistically invalid scheme $\exists x.[\exists x.\varphi \Rightarrow \varphi]$ provable), there seems to have been no attempt to investigate just how much extra strength ε confers on intuitionistic logic before [Bell 1993a]. And one of the things that he shows in that paper is that the addition of ε not only makes provable some strange quantifier rules, but that it also, given some almost trivial further assumptions, lets us prove rules that are often taken to be *laws of* *logic*, for instance one of de Morgan's laws.

This points up a reason why philosophers ought to take an interest in the theory of term forming operators. For the ε–operator can fairly naturally be regarded as an ontological principle of some sort, and yet de Morgan's law in some sense is a logical principle. So we have a derivation of a logical law from ontological assumptions. We will see in Chapter 11 that τ similarly allows us to derive principles of classical logic in intuitionistic logic. As Bell points out, such facts offer us the hope that we might be able to give a non-question begging justification of these laws of classical logic.

Probably the most important result in [Bell 1993a] is a proof that if we add both (Ack) and (ε) to intuitionistic logic, and if we make some almost trivial further assumptions, we can derive the law of excluded middle, and so we can derive all the laws of classical logic. On the other hand, Bell also shows that there are non-extensional models of the intuitionistic ε–calculus in which excluded middle fails. So the addition of the extensionality assumptions implicit in (Ack) are not harmless. So if we want a truly general theory of term forming operators we should expect it to be able to handle both extensional and non-extensional operators.

This puts the objections of Corcoran and Herring in a different light. If we accept that a general theory of term forming operators should be able to handle non-extensional operators, we must regard the failure of Universal Instantiation, the Substitutivity of Identicals, and so on, no longer as unacceptable problems but as consequences of the introduction of these non-extensional operators.

7. More on λ

The semantics for term forming operators in logic began to be developed almost immediately after the semantics for classical logic became a central part of logic. The semantics for the λ–calculus took much longer to develop. Indeed, while the λ–calculus had been fairly thoroughly researched from a proof-theoretic point of view by the early 1940's, there were no models, other than the so-called term models, which are syntactical in nature, before about 1970. Once again it was Dana Scott who was the first to develop these models, and he invented continuous lattices, now the subject of a considerable literature of their own, to do it. Since then there has been a revival of interest in λ–calculus, and it has been the subject of intense research.

Of particular interest to us is that in the model theory for the λ–calculus a central concern is the status of an extensionality principle which we might state, slipping in a quantifier and an implication symbol which properly speaking are part of the metalanguage,

$$\text{(w.e.)} \qquad \forall x.M = N \implies \lambda x.M = \lambda x.N.$$

It is interesting that it has been a subject of some debate whether it is legitimate to restrict our attention to models for λ which meet this extensionality condition in the name of naturalness and simplicity, even though it is not derivable from the rules of the λ–calculus.

8. Summary of What's to Come

This thesis contains a number of new results, notably the first completeness proofs for the ε– and τ–operators in first order intuitionistic logic, and a proposal for a framework for a general theory of term forming operators that is more flexible than any in the literature, since it handles non-extensional and intuitionistic cases. But I have attempted to write it so that it is accessible to non-experts since the community of experts on term forming operators is, even as communities of logicians go, pretty small. I have therefore attempted to write this thesis as a sort of an introductory text-book for the theory of term forming operators. Since it is destined to sit on a shelf in a philosophy department, I have tried to make it accessible to readers with a level of expertise that could be expected from someone with an intermediate course in logic in a philosophy department—a good understanding of completeness proofs in first order logic, some knowledge of rudimentary facts about lattices and sets, and some idea about the difference between intuitionistic and classical first-order logic.

Chapter 2 will be given over to setting out some of the mathematical definitions and theorems that will be appealed to in later chapters, both for the convenience of having them located in one place for ease of reference and to avoid having them interrupt the flow of later presentation.

Chapters 3 and 4 will be given over to a fairly self-contained presentation of the syntax and semantics, respectively, of the untyped λ–calculus.

In Chapter 5 we finally move on to the official topic of this thesis, namely term forming operators in first order logic. We give in that chapter a thorough presentation of the syntax for a first order language with a single variable binding term forming operator. Most authors simply wave their hands at some of the complications that result by adding a variable binding operator other than the quantifiers to a language, simply asserting that the appropriate modifications can be made to the syntactical machinery of the usual first order languages to ensure that things work out without a hitch. Authors who do give detailed syntactical definitions usually impose some restriction that is inappropriate for our purposes. For example, [Bencivenga et al.] allow only closed terms, which makes it much easier to give a definition of substitution; and [Leisenring] presents a syntax which is suitable for his purposes, since he gives only extensional semantics, but which is not suitable for non-extensional cases. The presentation in Chapter 5 is not restricted in these ways. That chapter also includes a presentation of the axioms of most of the systems we will deal with in subsequent chapters.

Chapter 6 is merely a collection of some of the facts about ordinary classical first order logic we will need to appeal to in later chapters.

Chapters 7, 8, and 9 are given over to a detailed discussion of the classical ε-calculus. Chapter 7 presents what is essentially the preferred semantics of [Asser], which was also the semantics adopted by [Leisenring], except that we generalize it so that it can handle open terms. We also prove a few other results, such as Hilbert's ε-theorems, using the soundness and completeness results for this semantics. In Chapter 8 we present a semantics for two different versions of the classical ε-calculus which do not satisfy (Ack). Chapter 9 presents the first investigation of a Boolean valued semantics for the ε-calculus. This will tell us something more about what needs to be done to get rid of the assumption of extensionality, but more importantly will tell us a lot about how to get rid of the assumption that there are only two truth-values.

Chapter 10 will again be nothing more than a collection of some facts about first order intuitionistic logic and its semantics for the sake of being able to appeal to them conveniently.

Chapter 11 will be an investigation of the intuitionistic ε-and τ-calculi.[7] We will present, among other things, the first complete semantics for two versions of each of the intuitionistic ε- and τ-calculi, making extensive use of what we learned in Chapters 8 and 9 about how to go about eliminating extensionality assumptions and the assumption that there are only two truth-values. This chapter will also include a variety of results, some new to the literature, about the extra deductive strength each of these operators adds to intuitionistic logic, and some independence results for the intuitionistic ε- and τ-calculi.

Finally, Chapter 12 will gather together various strands of our discussion, going all the way back to Chapter 3, to give a discussion of what earlier chapters teach us about the shape that a general theory of term forming operators in first order logic ought to take.

Before getting on with the main business of this thesis, perhaps something should

[7] Note that Chapters 7, 8, and 9 also implicitly cover the classical τ-calculus, since in classical logic ε and τ are interdefinable.

be said about what has not been included. First, no attempt has been made to develop a semantics for free logic. While it would have been possible to present a semantics for term forming operators in classical two-valued logic without greatly increasing the complexity of the presentation below, it seems that things are not so simple in the case of intuitionistic free logic. Such a semantics seems to require bringing in the heavy mathematical machinery of sheaf theory. This is a project for a later date. Since ι is only of very much semantic interest in a discussion of free logic, there is also almost no discussion of ι below.

We appealed above to the development of topos theory as one reason for the increased interest in intuitionistic logic. But this is really a reason for interest in *typed* intuitionistic logic. [Bell 1993b] presents a complete semantics for ε in typed intuitionistic logic. A general theory of term forming operators in this case would be of some interest, as would a comparison of the first order and typed cases, both of the respective general theories and of the theories for particular operators. As an example of the kind of question that would be worth pursuing, I mention that Bell is able to restrict attention to *closed* ε–terms in the typed case, since if we allow open ε–terms in typed intuitionistic logic we can derive the law of excluded middle and so our theory collapses to the classical case. In the first order case, intuitionistic logic with open ε–terms does not collapse into classical logic. What is behind this difference? An interesting question, but not one pursued here.

Finally, there has been some research, again mostly by da Costa and his collaborators, of the effect of term forming operators in "non-classical" logic other than intuitionistic logic. In particular, there has been some research on term forming operators in classical modal logic. We might say as a rough and ready approximation that we get non-classical logic in the form of intuitionistic logic by taking something out of classical logic, namely the law of excluded middle. We get classical modal logic by *adding* a modal operator to classical logic. Trying to include the modal logic case in our current treatment seemed to me to be going too far afield from our main focus. Those interested in pursuing that sort of study of term forming operators can find an overview in [da Costa and Mortensen], which also has references to relevant literature.

CHAPTER II

LEXICAL PRELIMINARIES

Before we go on, it will be useful to collect in one place many of the definitions and theorems that will be used in the remainder of this thesis, but which are probably already familiar to most readers. Removing them to one place will, I hope, make later chapters read more smoothly while still making the definitions available to those not familiar with them. In aid of readability many familiar or easily proved propositions and theorems will be stated without proof.

A *lattice* is a non-empty partially ordered set $L = \langle L, \leqslant \rangle$ in which each pair of elements $x, y \in L$ has a supremum $x \vee y \in L$ called the *join* of x and y and an infimum $x \wedge y \in L$ called the *meet* of x and y. We will often identify a lattice with its underlying set when confusion is unlikely.

A lattice L is *distributive* if for all $x, y, z \in L$,

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

If L is distributive, then for all $x, y, z \in L$, $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$.

Obviously, in any lattice L every finite non-empty subset of L has a meet and a join. If *every* (finite or infinite) subset of L has a meet and a join, we say that L is a *complete lattice*. (So every finite lattice is complete.) For arbitrary $X \subseteq L$, we denote the join of X, if there is one, by '$\bigvee X$', and the meet of X, if there is one, by '$\bigwedge X$'. If $X = \{ x_i \mid i \in I \}$ is an indexed set, we will sometimes write $\bigvee_{i \in I} x_i$ and $\bigwedge_{i \in I} x_i$ for the join and meet, if they exist.

We will write 0 and 1 for the *least* and *greatest* elements of L, if they exist. Suppose L has 0 and 1. An element $y \in L$ is a *complement* for $x \in L$ if both $x \vee y = 1$ and $x \wedge y = 0$. If every element of L has a complement, we say that L is *complemented*.

Proposition 1. *An element of a distributive lattice has at most one complement.* □

Obviously, then, in a complemented distributive lattice we can define a map $x \mapsto x^*$ taking each element x to its unique complement x^*.

Definition. A *Boolean algebra* is a complemented distributive lattice.

We will assume that in each Boolean algebra $0 \neq 1$. A non-empty subset B' of a Boolean algebra B which is closed under $\wedge$, $\vee$ and * is a (Boolean) *subalgebra* of B.

We define the notion of relative pseudo-complementation in a lattice as follows. For $a, b \in L$, the *relative pseudo-complement of* a *to* b, $a \Rightarrow b$, if it exists, is

$$\bigvee \{ x \mid x \wedge a \leqslant b \}.$$

If L is a Boolean algebra, $a \Rightarrow b = a^* \vee b$, and so always exists. A lattice L in which $a \Rightarrow b$ exists for all $a, b \in L$ is said to be *relatively pseudo-complemented*.

Definition. A *Heyting algebra* is a relatively pseudo-complemented lattice with 0.

It is easy to show that Heyting algebras are distributive. In a Heyting algebra $a \Rightarrow 0$ always exists, and is called the *pseudo-complement* of a. Each Boolean algebra is a Heyting algebra, and $a \Rightarrow 0 = a^*$. We therefore extend the a^* notation to denote the pseudo-complement of a in any Heyting algebra.

A *filter* in a Heyting algebra H (and so in a Boolean algebra B) is a non-empty subset $\mathbf{F}$ of L such that for $x, y \in H$,

(1) $x, y \in \mathbf{F} \implies x \wedge y \in \mathbf{F}$

(2) $x \in \mathbf{F}$ and $x \leqslant y \implies y \in \mathbf{F}$

(3) $0 \notin \mathbf{F}$.

A *Boolean algebra homomorphism* is a map $h : B \to B'$, with B and B' both Boolean algebras, such that for $x, y \in B$,

$$h(x \wedge y) = h(x) \wedge h(y)$$

$$h(x \vee y) = h(x) \vee h(y)$$

$$h(x^*) = h(x)^*$$

where the operations on the left are in B while those on the right are in B'.

A *Heyting algebra homomorphism* is a map $h : H \to H'$, with H, H' both Heyting algebras, such that, for $x, y \in H$,

$$h(x \wedge y) = h(x) \wedge h(y)$$

$$h(x \vee y) = h(x) \vee h(y)$$

$$h(x \Rightarrow y) = h(x) \Rightarrow h(y)$$

where the operations on the left are in H, while those on the right are in H'. Clearly, since $a \Rightarrow b = a^* \vee b$ in a Boolean algebra, each Boolean algebra homomorphism is also a Heyting algebra homomorphism. In either case, if h is one-one and onto, we say h is an *isomorphism* (of the appropriate sort), and write '$H \cong H'$' ('$B \cong B'$').

Proposition 2. *If $h : H \to H'$ is a homomorphism, then $h^{-1}[1] = \{ x \in H \mid h(x) = 1 \}$ is a filter in H.* □

This filter is called the *hull* of h. For each filter $\mathbf{F}$ of H, we define the *quotient algebra* as follows. First, define for $a, b \in H$

$$x \Leftrightarrow y = (x \Rightarrow y) \wedge (y \Rightarrow x).$$

Next, define the relation $\sim_{\mathbf{F}}$ on H by putting

$$x \sim_{\mathbf{F}} y \iff x \Leftrightarrow y \in \mathbf{F}.$$

This is easily shown to be a congruence relation, so if we put, for each $x \in H$,

$$[x] = \{ y \in H \mid y \sim_{\mathbf{F}} x \}$$

we can define a Heyting algebra $H/\mathbf{F}$ by putting

$$[x] \vee [y] = [x \vee y]$$

$$[x] \wedge [y] = [x \wedge y]$$

$$[x] \Rightarrow [y] = [x \Rightarrow y].$$

Proposition 3. (1) *The map $h : H \to H/\mathbf{F}$ defined by $h(x) = [x]$ is a homomorphism of H onto $H/\mathbf{F}$. If H is a Boolean algebra, then $H/\mathbf{F}$ is a Boolean algebra* *and h is a Boolean algebra homomorphism.* $\mathbf{F}$ *is the hull of h.* (2) *Let $h : H \to H'$ be a homomorphism, and let* $\mathbf{F}$ *be the hull of h. Then $h[H] \cong H/\mathbf{F}$.* □

The map $x \mapsto [x]$ is called the *canonical homomorphism*.

An *ultrafilter* in a lattice L is a filter $\mathbf{F}$ such that for any filter $\mathbf{F}'$,

$$\mathbf{F} \subseteq \mathbf{F}' \implies \mathbf{F} = \mathbf{F}'.$$

A *prime filter* in L is a filter $\mathbf{F}$ such that

$$x \vee y \in \mathbf{F} \implies x \in \mathbf{F} \text{ or } y \in \mathbf{F}.$$

It is easy enough to verify that every two element lattice is isomorphic to the lattice $\mathbf{2} = \langle \{0,1\}, \leqslant \rangle$, where $\leqslant$ is the relation that puts $0 \leqslant 0$, $0 \leqslant 1$ and $1 \leqslant 1$, and that $\mathbf{2}$ is in fact a Boolean algebra. A homomorphism $h : H \to \mathbf{2}$ is called a $\mathbf{2}$–valued homomorphism.

Proposition 4. *Let* $\mathbf{F}$ *be a filter in a Heyting algebra H. Then the following conditions are equivalent.*

(1) $H/\mathbf{F} \cong \mathbf{2}$

(2) $\mathbf{F}$ *is the hull of a* $\mathbf{2}$*–valued homomorphism on H*

(3) $\mathbf{F}$ *is an ultrafilter*

(4) *For each $x \in H$, either $x \in \mathbf{F}$ or $x^* \in \mathbf{F}$.*

If H is a Boolean algebra, we can add to the list

(5) $\mathbf{F}$ *is a prime filter.* □

It is easily shown that the union of any chain of filters in a lattice L is itself a filter, and that the union would be an upper bound in the set of filters of L ordered by inclusion. So Zorn's Lemma can be invoked to get the

Theorem 1 (Ultrafilter Theorem). *Each filter in a lattice is included in an ultrafilter.* □

A subset X of a lattice L has the *finite meet property* (fmp) if

$$x_1, \ldots, x_n \in X \implies x_1 \wedge \cdots \wedge x_n \neq 0.$$

An easy corollary of the Ultrafilter Theorem is

Corollary. *Let $X \subseteq L$. Then X has the fmp $\iff$ X is included in an ultrafilter.* □

We pause here to note that while Zorn's Lemma entails the Ultrafilter Theorem, the converse is not the case. As we shall see below, if we are willing to content ourselves with weaker completeness theorems we can sometimes simply adopt the ultrafilter theorem as a basic principle, but if we want strong completeness theorems, we will need the full strength of Zorn's Lemma (which is well known to be equivalent to the Axiom of Choice).

One important fact about the Ultrafilter Theorem is that it allows us to prove the Rasiowa–Sikorski Theorem, which will be of interest to us later. Let H be a Heyting algebra, U an ultrafilter in H, and $T \subseteq H$ such that $\bigvee T$ exists. We say that U *respects* T (or $\bigvee T$) if

$$\bigvee T \in U \implies T \cap U \neq \emptyset.$$

Let $\mathcal{T}$ be a family of subsets of H such that each member of $\mathcal{T}$ has a join in H. Then U respects $\mathcal{T}$ (or $\{\bigvee T \mid T \in \mathcal{T}\}$) if U respects each member of $\mathcal{T}$.

It is an interesting and unfortunate fact that there are families of subsets of complete Boolean algebras (and so of Heyting algebras) which are respected by no ultrafilter (see [Bell and Machover, p. 158] for an example). The next theorem shows that things are quite different if we restrict our attention to countable families.

Theorem 2 (Rasiowa–Sikorski Theorem). *Let H be a Heyting algebra, and $\mathcal{T}$ a countable family of subsets of H such that $\bigvee T$ exists for all $T \in \mathcal{T}$. Then there is an ultrafilter U in H which respects $\mathcal{T}$.* $\square$

We will also want to have available the notion of a finitary closure operation in what follows.

Let S be a set. A *finitary closure operation on S* is a map $\mathbf{C} : \mathcal{P}(S) \to \mathcal{P}(S)$ such that for $X, Y \subseteq S$

(1) $X \subseteq \mathbf{C}(X)$

(2) $\mathbf{C}(\mathbf{C}(X)) \subseteq \mathbf{C}(X)$

(3) $Y \subseteq X \implies \mathbf{C}(Y) \subseteq \mathbf{C}(X)$

(4) $x \in \mathbf{C}(X) \implies$ there is a finite Y, $Y \subseteq X$ and $x \in \mathbf{C}(Y)$.

It will be useful to have the following definition. Let $\mathbf{f}$ be a distinguished element of S. We say that X is *maximal under* $\mathbf{C}$, and write '$Max_{\mathbf{C}}(X)$', if

(1) $\mathbf{f} \notin \mathbf{C}(X)$

(2) $X \subseteq Y \subseteq S$ and $\mathbf{f} \notin \mathbf{C}(Y) \implies X = Y$.

Obviously if $\mathbf{C}$ is a finitary closure operation on S, $Max_{\mathbf{C}}(X) \implies \mathbf{C}(X) = X$.

Theorem 3. *If* $\mathbf{f} \notin \mathbf{C}(X)$*, then there is a* Y *such that* $X \subseteq Y \subseteq S$ *and* $Max_{\mathbf{C}}(Y)$.

Proof. Consider the set $\mathcal{Z} = \{ Z \subseteq S \mid X \subseteq Z \subseteq S \text{ and } \mathbf{f} \notin Z \}$, ordered by inclusion. If $\mathcal{F}$ is a chain in this set, it is easily shown that $\bigcup_{F \in \mathcal{F}} F \in \mathcal{Z}$. So the result follows by Zorn's Lemma. □

In our chapter on semantics for the λ-calculus we will need a very few concepts from category theory.

A *category* $\mathbf{C}$ consists of two classes $\mathrm{Ob}(\mathbf{C})$ (the objects of $\mathbf{C}$) and $\mathrm{Arr}(\mathbf{C})$ (the arrows of $\mathbf{C}$) and two maps $\mathrm{cod}, \mathrm{dom} : \mathrm{Arr}(\mathbf{C}) \to \mathrm{Ob}(\mathbf{C})$ which satisfy (1)–(4) listed below. If $f \in \mathrm{Arr}(\mathbf{C})$, $\mathrm{dom}(f) = A$, and $\mathrm{cod}(f) = B$ we say that f is an *arrow from* A *to* B, and write $f : A \to B$.

(1) For any pair of arrows $f, g \in \mathrm{Arr}(\mathbf{C})$ such that $\mathrm{cod}(f) = \mathrm{dom}(g)$, $f : X \to Y$, $g : Y \to Z$, there is an arrow $g \circ f \in \mathrm{Arr}(\mathbf{C})$, called the *composition of* f *and* g such that $g \circ f : X \to Z$.

(2) For $X \in \mathrm{Ob}(\mathbf{C})$, there is an *identity arrow* $1_X : X \to X$.

(3) Compositions satisfy the associative law, viz. for arrows $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$ of $\mathbf{C}$, $h \circ (g \circ f) = (h \circ g) \circ f$.

(4) For any object Y of $\mathbf{C}$ and arrows $f : X \to Y$, $g : Y \to Z$ of $\mathbf{C}$, $1_Y \circ f = f$ and $g \circ 1_Y = g$.

We say that two objects A and B are *isomorphic* (and we write '$A \cong B$') if there are arrows $f : A \to B$ and $g : B \to A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$.

A *terminal object* $\mathbf{1} \in \mathrm{Ob}(\mathbf{C})$ is an object such that for every $A \in \mathrm{Ob}(\mathbf{C})$ there is a *unique* $f \in \mathrm{Arr}(\mathbf{C})$ such that $f : A \to \mathbf{1}$. For $A_1, A_2 \in \mathrm{Ob}(\mathbf{C})$ a *cartesian product*, is an object $A_1 \times A_2 \in \mathrm{Ob}(\mathbf{C})$ together with a pair of (projection) arrows $\pi_1, \pi_2 \in \mathrm{Arr}(\mathbf{C})$ with $\pi_1 : A_1 \times A_2 \to A_1$, $\pi_2 : A_1 \times A_2 \to A_2$ such that for any pair of arrows $f_1 : C \to A_1$, $f_2 : C \to A_2$ of $\mathbf{C}$ there is a *unique* arrow $\langle f_1, f_2 \rangle : C \to A_1 \times A_2$ such that $\pi_1 \circ \langle f_1, f_2 \rangle = f_1$ and $\pi_2 \circ \langle f_1, f_2 \rangle = f_2$.

In a category there can be any number of terminal objects and, for any pair of objects, any number of cartesian products. However, in a particular category any two terminal objects and any pair of products for a given pair of objects are isomorphic. For convenience we will assume that we have selected one terminal object and one product for each pair of objects, provided in each case that there is at least one. Assuming that the required products exist, if $g_1 : A_1 \to B_1$ and $g_2 : A_2 \to B_2$, we will put $g_1 \times g_2 = \langle g_1 \circ \pi_1, g_2 \circ \pi_2 \rangle : A_1 \times A_2 \to B_1 \times B_2$.

We say that a category $\mathbf{C}$ is a *cartesian closed category (ccc)* if and only if (1) $\mathbf{C}$ has a terminal object $\mathbf{1}$; (2) Every pair of objects in $\mathbf{C}$ has a cartesian product; (3) For $A, B \in \mathrm{Ob}(\mathbf{C})$ there is an *exponential*, i. e. an object $B^A \in \mathrm{Ob}(\mathbf{C})$ together with a map $ev_{A,B} : B^A \times A \to B$ such that for all $f : C \times A \to B$ there is a *unique* arrow $\bar{f} : C \to B^A$ (called the *transpose* of f) satisfying $f = ev_{A,B} \circ \bar{f} \times 1_A$. It is easy to show that there is a bijective correspondence between arrows from $C \times A \to B$ and their transposes.

It will be useful to have the following definitions.

We will write $hom_{\mathbf{C}}(A, B)$ (or $hom(A, B)$ if the category is obvious) for the class of $f \in \mathrm{Arr}(\mathbf{C})$ with $\mathrm{dom}(f) = A$ and $\mathrm{cod}(f) = B$.

If $\mathbf{C}$ is a ccc, then for any object $A \in \mathrm{Ob}(\mathbf{C})$, define $A^0 = \mathbf{1}$, $A^{n+1} = A^n \times A$. It is not hard to show that $A \cong A^1$.

Let $\Delta = x_1, \ldots, x_n$ be a sequence of distinct variables. We then sometimes write 'A^Δ' for A^n. We write '$\pi^\Delta_{x_i} : U^\Delta \to U$' for the projection on the i-th coordinate.

For n arrows of $\mathbf{C}$, $f_1, \ldots, f_n : A \to U$, we define $\langle f_1, \ldots, f_n \rangle : A \to U^n$ inductively by

$$\langle\rangle = f : A \to \mathbf{1}$$

(This, recall, is unique since $\mathbf{1}$ is terminal.)

$$\langle f_1, \ldots, f_{n+1} \rangle = \langle \langle f_1, \ldots, f_n \rangle, f_{n+1} \rangle .$$

Clearly $\pi^\Delta_{x_i} \circ \langle f_1, \ldots, f_n \rangle = f_i$.

Let $\Gamma = y_1, \ldots, y_m$ be a sequence of variables with each $y_i = x_j$ for some x_j of Δ, and where the ordering of Γ is the restriction of the ordering of Δ. Define $\Pi^\Delta_\Gamma = \langle \pi^\Delta_{y_1}, \ldots, \pi^\Delta_{y_m} \rangle : U^\Delta \to U^\Gamma$. We will abuse notation in the following ways. We will write '$\{\Gamma\}$' for the set of elements in the sequence Γ, but will write '$\{\Gamma\} \subseteq \{\Delta\}$' if Γ is a subsequence of Δ in the way described at the start of this paragraph.

One notion that will be especially important later on is the notion of a category having *enough points*. Let $\mathbf{C}$ be a category with terminal object $\mathbf{1}$. A *point* of $A \in \mathrm{Ob}(\mathbf{C})$ is a $\mathbf{C}$-arrow $x : \mathbf{1} \to A$. The class of points of A is denoted by $|A|$. A has *enough points* if for any $f, g : A \to A$

$$f \neq g \implies f \circ x \neq g \circ x$$

for some $x \in |A|$.

In a ccc there is a bijective correspondence between arrows $A \to B$ and points of B^A. If $f : B \to A$, we will write '' f'' for the corresponding point of B^A and call this the *name* of f.

Let **C** be a ccc. An object $U \in \mathrm{Ob}(\mathbf{C})$ is *reflexive* if U^U is a retract of U, that is, if there are arrows $F : U \to U^U$ and $G : U^U \to U$ such that $F \circ G = 1_{U^U}$.

We will make extensive use of the following proposition in Chapter 4.

Proposition 5. *In a ccc*

(1) $\overline{h \circ g \times 1_A} = \bar{h} \circ g$

(2) $\langle f, g\rangle \circ h = \langle f \circ h, g \circ h\rangle$

(3) $f \times g \circ \langle h, k\rangle = \langle f \circ h, g \circ k\rangle$. $\square$

Finally, to state some of our results about different sorts of models for the λ-calculus we will need the following notions. Let **C** and **D** be categories. A *functor* $F : \mathbf{C} \to \mathbf{D}$ is a function taking $\mathrm{Ob}(\mathbf{C})$ to $\mathrm{Ob}(\mathbf{D})$ and $\mathrm{Arr}(\mathbf{C})$ to $\mathrm{Arr}(\mathbf{D})$ such that: if $f : X \to Y$ in **C**, then $F(f) : F(X) \to F(Y)$; if $g \circ f \in \mathrm{Arr}(\mathbf{C})$, then $F(g \circ f) = F(g) \circ F(f)$; and $F(1_X) = 1_{F(X)}$ for all $X \in \mathrm{Ob}(\mathbf{C})$.

Let $F : \mathbf{C} \to \mathbf{D}$ be a functor. F is *full* if for $A, B \in \mathrm{Ob}(\mathbf{C})$ F takes $hom_{\mathbf{C}}(A, B)$ *onto* $hom_{\mathbf{D}}(F(A), F(B))$. F is *faithful* if for $A, B \in \mathrm{Ob}(\mathbf{C})$, F is one-to-one on $hom_{\mathbf{C}}(A, B)$. F is *dense* if for each $B \in \mathrm{Ob}(\mathbf{D})$ there is an $A \in \mathrm{Ob}(\mathbf{C})$ such that $F(A) = B$.

If $F : \mathbf{C} \to \mathbf{D}$ is a bijective functor between two categories **C** and **D**, we say that F is an *isomorphism* and that **C** and **D** are *isomorphic categories*. A weaker but more common notion is equivalence of categories. We say that F is an *equivalence* and that $\mathbf{C}$ and $\mathbf{D}$ are *equivalent* categories if F is full, faithful, and dense.

Readers interested in a more detailed but still introductory account of Boolean algebras are directed to [Bell and Machover, ch. 4]. A useful introduction to lattice theory, Boolean algebras, and Heyting algebras aimed at logicians can be found in [Rasiowa and Sikorski, ch. 1–4]. (They use the term "pseudo–Boolean algebra" for what we have called a Heyting algebra.) A more comprehensive treatment can be found in [Balbes and Dwinger]. Those interested in finitary closure operations and their relationship to logic would do well to begin at the beginning with [McKinsey and Tarski 1944, 1946], though the operators considered there are slightly different from those used here. Useful introductions to category theory for logicians which include the rudimentary bits we will use here can be found in the early chapters of [McLarty] and in chapter 1 of [Bell, 1988].

CHAPTER III

λ——CALCULUS: SYNTACTICAL INVESTIGATIONS

Perhaps the most familiar and certainly the most studied term forming operator is the abstraction operator λ in Church's λ–calculus. We mentioned above that our focus in this thesis will be term forming operators which bind a single variable in first order predicate logic. As we shall see λ, strictly speaking, falls outside this range because if we restrict the λ–calculus to the *untyped* case and try to include negation or implication, we get inconsistency. We therefore cannot treat λ as a term forming operator in standard first-order logical systems. Nonetheless, it is useful to our project to review some of the features of the λ–calculus. First, precisely because the λ–calculus is so well understood it is important to make clear the reason that the model theory for λ cannot be carried over to other operators. More positively, λ is a term forming operator that binds variables, and such things normally do not appear in treatments of first order logic. We therefore can expect some hints about how to design the syntactical machinery in first order logic from the way these problems are handled in the λ case. It also turns out that some of the problems which arise in trying to design models for the λ–calculus have analogues in the case of term forming operators in first order logic. Finally, some of the properties and theorems which have their home in the λ–calculus (such as the Church–Rosser property) will play an important role in our general semantics for term forming operators, even if it is a very different role from the one they play in the λ-calculus. Since our interest in the λ-calculus is limited, the presentation in the next two chapters will be introductory but rather selective. Those interested in a full, detailed treatment of the λ-calculus are directed to the thorough [Barendregt 1981, revised edition 1984] while those interested particularly in the model theory are directed to the shorter work [Koymans]. A good introductory presentation, though one with a rather different emphasis from the one here, can be found in [Hindley and Seldin].

1. λ–Terms

The usual inductive definition of the terms of a language $\mathcal{L}$ for first order logic is well known and simple: the variables and constants of $\mathcal{L}$ are terms; if $t_1, \ldots, t_n$ are terms of $\mathcal{L}$ and f is an n-ary function symbol of $\mathcal{L}$, then $f(t_1, \ldots, t_n)$ is a term of $\mathcal{L}$. Once a person gets a handle on this definition in a first course in logic, there is normally little reason ever to think of it again. Things are not so simple when we allow term forming operators such as ε or τ into our system, since these operators bind a variable. The result is that the syntactical problems of variable collision, which we can usually put off until we set out the quantifier rules, must be dealt with much sooner.

Fortunately, a lot of work has already been done on the syntax of a term forming operator which binds a variable, namely Church's abstraction operator λ. Since the λ-calculus is a *term calculus*, that is, there is no distinction between the terms and formulas in the λ-calculus, the abstraction operator gives us a term of the language

$(\lambda x.M)$ for each *term* M of the language, while ε, for example, gives us a term $\varepsilon x.\varphi$ for each *formula* φ, where x is a variable of the respective languages. Nonetheless we will be able to transfer many ideas encountered in the λ case directly to the case of first order logic in chapter 5. We therefore develop some of these points in what most readers will no doubt find excruciating detail.

We assume given a countably infinite set of variables $\mathbf{Var} = \{v_1, v_2, v_3, \dots\}$, and a set, which might be infinite, finite, or empty, of constants $\mathbf{C} = \{c_1, c_2, c_3, \dots\}$.

Definition 1. The set $\Lambda(\mathbf{C})$ of *λ-terms with constants from* $\mathbf{C}$ is the smallest set such that

(1) $\mathbf{Var} \subseteq \Lambda(\mathbf{C})$

(2) $\mathbf{C} \subseteq \Lambda(\mathbf{C})$

(3) $M, N \in \Lambda(\mathbf{C}) \implies M(N) \in \Lambda(\mathbf{C})$

(4) $M \in \Lambda(\mathbf{C}), x \in \mathbf{Var} \implies (\lambda x.M) \in \Lambda(\mathbf{C})$.

If $\mathbf{C}$ is empty, the resulting system is called the *pure λ-calculus*; otherwise it is called *applied*. The variables are called *atoms*. A term of form $M(N)$ is called an *application of M to N*. Terms of form $(\lambda x.M)$ are called *abstraction terms*.

Notation. We will use capital Roman letters to denote arbitrary λ-terms. We use the letters 'x', 'y', 'z', 'u', 'v', 'w' to denote arbitrary variables, and assume that distinct letters denote distinct variables unless the contrary is explicitly stated. We will also occasionally use brackets in place of parentheses to enhance the readability of λ-terms when such usage is unlikely to cause confusion with our notation for substitution.

Examples of λ-terms in pure λ-calculus.

$$(\lambda x.x)(y) \qquad (\lambda x.x)((\lambda y.y)(x))$$

$$(\lambda x.y)(x) \qquad x(\lambda x.(\lambda x.x))$$

Informally speaking, the terms can be read as follows: where M has been interpreted as a map, $M(N)$ can be regarded as the result of applying M to the argument N. The usual notation for this in the literature on the λ-calculus and combinatory logic is (MN), but for readability among non-specialists we will use the usual notation for function application. The term $(\lambda x.M)$ is intended to represent the map whose value at an argument N is calculated by substituting N for x in M.

2. Free and Bound Variables, Substitution, α–Convertibility, and Syntactical Congruence

As we have already mentioned, the abstraction operator binds variables in much the same way as do the quantifiers of predicate logic. That is, if M is a λ-term (in which x may or may not occur free), x does not occur free in $(\lambda x.M)$. Not surprisingly, then, some care needs to be taken when spelling out the conditions under which it is permissible to substitute one term for another within a λ-term. Substitution is going to be important to us, as a look at the intuitive meaning of the abstraction terms mentioned above makes clear.

There is an important choice which must be made in any presentation of the λ-calculus. If we (temporarily) use '≡' as a metatheoretical symbol to indicate syntactical equality, and write $M[x/y]$ for the result of replacing all free occurrences of x in M by y, then, provided y is not free in M, we shall clearly want to do something to ensure that the difference between $(\lambda x.M)$ and $(\lambda y.M[x/y])$ turns out to be insignificant. In Church's formulation of the λ-calculus this was achieved by including as a basic axiom of the λ-calculus the

(α-axiom) $\quad (\lambda x.M) = (\lambda y.M[x/y])$ $\quad$ (provided y does not occur free in M).

The main disadvantage of this is that it makes some of the theorems rather cumbersome to write because it necessitates regularly spelling out some not very interesting conditions on allowable substitutions.

The other approach is to define a notion of α-convertibility, which we write '$\equiv_\alpha$', by saying that two λ-terms are α-convertible if they can each be converted to the other by a series of renamings of bound variables. We then adopt the convention that if $M \equiv_\alpha N$, we simply say that $M \equiv N$. We also adopt the convention that when we deal with $M_1, \ldots, M_n$ in a particular context (e.g. a proof) we assume that all the bound variables are chosen to be different from the free variables. This allows us to work on λ-terms without worrying about whether substitutions are permissible or not. On the other hand, if we adopt this approach the λ-terms we work with are merely representatives of α-convertibility-equivalence classes. The result is that we need to take extra care in our definition of substitution, and to pause to show that operations on terms determine well-defined operations on $\equiv_\alpha$-equivalence classes. The proof of this is so long and tedious that it is not even reproduced in the encyclopedic [Barendregt 1981]. Nonetheless, we shall adopt this second approach and will direct readers to [Curry and Feys] for the more tedious details of some of the proofs.

Notation. We denote *syntactical identity* of λ-terms by '$\equiv$'; so $M \equiv N$ means that M is exactly the same term as N.

Definition 2. We define P *is a subterm of* Q (or P *occurs in* Q) inductively by:

(1) P is a subterm of P

(2) P is a subterm of M or of $N \implies P$ is a subterm of $M(N)$

(3) P is a subterm of M or $P \equiv x \implies P$ is a subterm of $(\lambda x.M)$.

For an occurrence of $(\lambda x.M)$ in a term P, that occurrence of M is called the *scope* of that occurrence of λ. An occurrence of a variable x in a term P is called *bound* if and only if it is in a subterm of P of the form $(\lambda x.M)$; otherwise it is called *free*.

Definition 3. The *set of free variables of* M, denoted by '$\mathbf{FV}(M)$', is defined inductively by:

(1) $\mathbf{FV}(x) = \{x\}$

(2) $\mathbf{FV}(\lambda x.M) = \mathbf{FV}(M) - \{x\}$

(3) $\mathbf{FV}(M(N)) = \mathbf{FV}(M) \cup \mathbf{FV}(N)$.

It is easily checked that the set of free variables for a term P is precisely the set of variables which have a free occurrence in P.

Definition 4. Let $M, N \in \Lambda(\mathbf{C})$. Then *the result of substituting N for x in M*, denoted by '$M[x/N]$', is defined inductively as follows:

(1) $M \equiv x \implies M[x/N] \equiv N$

(2) $M \equiv y \not\equiv x \implies M[x/N] \equiv y$

(3) $M \equiv M_1(M_2) \implies M[x/N] \equiv M_1[x/N](M_2[x/N])$

(4) $M \equiv (\lambda x.M_1) \implies M[x/N] \equiv (\lambda x.M_1)$

(5) $M \equiv (\lambda y.M_1)$ and $x \not\equiv y \implies M[x/N] \equiv (\lambda z.M_1[y/z][x/N])$

where in (5) $z \equiv y$ if $x \notin \mathbf{FV}(M_1)$ or $y \notin \mathbf{FV}(N)$, otherwise z is the first variable in the sequence $v_1, v_2, v_3, \ldots$ which occurs neither in M_1 nor in N.

A *change of bound variables* in M is the replacement of a subterm $(\lambda x.N)$ of M by $(\lambda y.N[x/y])$, where y does not occur in N. M is said to be *α-congruent* with N, and we write '$M \equiv_\alpha N$', if N can be obtained from M by a finite (possibly empty) series of replacements of bound variables.

Variable Convention 1. We identify α-congruent terms; that is, we redefine $M \equiv N$ to mean $M \equiv_\alpha N$.

In order to show that this convention doesn't cause us any problems we need the following

Proposition 1.

(1) $M \equiv_\alpha M', N \equiv_\alpha N' \implies M(N) \equiv_\alpha M'(N')$

(2) $M \equiv_\alpha M' \implies \lambda x.M \equiv_\alpha \lambda x.M'$

(3) $M \equiv_\alpha M', N \equiv_\alpha N' \implies M[x/N] \equiv_\alpha M'[x/N']$. $\square$

For a proof, see [Curry and Feys, pp. 94–104].

Variable Convention 2. If in a particular definition, proof, or other context the terms $M_1, \ldots, M_n$ occur, then in these terms all bound variables are chosen to be different from the free variables.

This convention is obviously legitimate whenever n is finite. In all such contexts, then, we can work with λ-terms as representatives of $\equiv_\alpha$-congruence classes and can perform substitutions without worrying about their permissibility.

It will also be useful to have available the notion of a simultaneous substitution. The conventions just mentioned make the definition fairly straightforward.

Definition 5. Let $M \in \Lambda(\mathbf{C})$, $\vec{x} = \langle x_1, \ldots, x_n \rangle$ be a sequence of distinct variables, and $\vec{N} = \langle N_1, \ldots, N_n \rangle$ a sequence of terms. The *simultaneous substitution of $\vec{N}$ for $\vec{x}$ in M*, which is written either as '$M[x_1/N_1, \ldots, x_n/N_n]$' or as '$M[\vec{x}/\vec{N}]$', is defined inductively as follows:

(1) $M \equiv y \equiv x_i \in \{x_1, \ldots, x_n\} \implies M[\vec{x}/\vec{N}] \equiv N_i$

(2) $M \equiv y \notin \{x_1, \ldots, x_n\} \implies M[\vec{x}/\vec{N}] \equiv y$

(3) $M \equiv M_1(M_2) \implies M[\vec{x}/\vec{N}] \equiv M_1[\vec{x}/\vec{N}](M_2[\vec{x}/\vec{N}])$

(4) $M \equiv (\lambda y.M_1) \implies M[\vec{x}/\vec{N}] \equiv (\lambda y.M_1[\vec{x}/\vec{N}])$.

It is straightforward to show that this does what we want in cases where some of the x_i occur free in some of the N_j, and no renaming of bound variables is needed, by Variable Convention 2. We should keep in mind, though, that $M[x_1/N_1, \ldots, x_n/N_n]$ can differ from $M[x_1/N_1] \ldots [x_n/N_n]$.

3. λ–Theories and Logic

We are reviewing the λ-calculus not merely in order to get hints for the syntactic features of our treatment of other term forming operators, but also because the model theory for the λ-calculus can offer us some hints about how to handle some of the problems that will confront us when designing models for other term forming operators below. Before we move on to talk about the models of the λ-calculus, though, we need to know what counts as a λ-theory. It will also be useful to review some of the key facts about these theories because this will help in the development of the model theory for λ–calculus.

However, there is a still more important reason for devoting as much effort as we are about to do to giving a detailed proof of the Church–Rosser Theorem and some of its corollaries, even though they are well-known results and have been for almost 60 years. It is the *details* of this proof that we will be appealing to in Chapter 12, so the details would need to be presented either here or there (since the details here are presented in a form quite different from that in the usual introductions to λ–calculus.) We will include them here because this is, so to speak, their natural home.

3.1 λ–Theories and the Church–Rosser Theorem.

For $M \in \Lambda(\mathbf{C})$, if $\mathbf{FV}(M) = \emptyset$, then we say that M is *closed*. If a closed term N contains no constants, then we sometimes say that N is a *combinator*. We will define $\Lambda^0(\mathbf{C}) \stackrel{\text{def}}{=} \{\, M \in \Lambda(\mathbf{C}) \mid M \text{ is closed} \,\}$. If $M, N \in \Lambda(\mathbf{C})$, then $M = N$ is a *formula* (or *equation*) of $\Lambda(\mathbf{C})$.[1] If both M and N are closed, $M = N$ is a *sentence* (or a *closed formula*, or a *closed equation*) of $\Lambda(\mathbf{C})$.

Definition 6. The *axioms of the λ-calculus* are all the instances of the following schemata[2]:

(β-schema) $$(\lambda x.M)(N) = M[x/N]$$

(ρ-schema) $$M = M.$$

We adopt the following as *rules of inference for the λ-calculus*[3]:

(σ) $$\frac{M = N}{N = M}$$

(τ) $$\frac{M = N \qquad N = P}{M = P}$$

(μ) $$\frac{M = M'}{N(M) = N(M')}$$

(ν) $$\frac{M = M'}{M(N) = M'(N)}$$

(ξ) $$\frac{M = M'}{(\lambda x.M) = (\lambda x.M')}$$

We will need a definition of a proof in this calculus, but any of the usual definitions will do. So we define a *derivation* of a formula F *from a set of formulas* S to be a sequence of formulas, the last of which is F, and all of which are either axioms or members of S, or which follow from preceding members of the sequence by one of the rules of inference.

[1] Note that we are using the same symbol '=' in formulas and in the metalanguage. Context will always make clear in which way the symbol is being used in any particular case.

[2] We should probably note that if we had taken the other approach to presenting the λ-calculus we would have needed to include the α-axiom here.

[3] The rule names are the standard ones from [Curry and Feys].

Let $\mathbf{T}$ be a set of $\Lambda(\mathbf{C})$ equations. '$\mathbf{T} \vdash M = N$' means there is a derivation of $M = N$ from $\mathbf{T}$. $\mathbf{T}$ is a $\Lambda(\mathbf{C})$ *theory* if $\mathbf{T}$ is a set of *closed* formulas such that $\mathbf{T} \vdash M = N \implies (M = N) \in \mathbf{T}$ for all *closed* $M, N \in \Lambda(\mathbf{C})$. We write '$\vdash M = N$' for '$\emptyset \vdash M = N$'.

Our goal now is to prove the most famous result about the λ-calculus, namely the Church-Rosser Theorem. Before we can even express this theorem, we will need to introduce another formal theory, namely the theory of λ-reduction. (This theory is also often called the theory of β-reduction). First we need some more definitions.

The set of terms of the theory of λ-reduction is $\Lambda(\mathbf{C})$. If $M, N \in \Lambda(\mathbf{C})$, then $M \triangleright N$ is a *formula* of $\Lambda(\mathbf{C})$. If both M and N are closed, $M \triangleright N$ is a *sentence* (or a *closed formula*) of $\Lambda(\mathbf{C})$.

Definition 7. The *axioms for the theory of λ-reduction* are all the instances of the following schemes:

$$(\triangleright\beta) \qquad (\lambda x.M)(N) \triangleright M[x/N]$$

$$(\triangleright\rho) \qquad M \triangleright M.$$

We adopt the following as *rules of inference for the theory of λ-reduction*:

$$(\triangleright\tau) \qquad \frac{M \triangleright N \qquad N \triangleright P}{M \triangleright P}$$

$$(\triangleright\mu) \qquad \frac{M \triangleright M'}{N(M) \triangleright N(M')}$$

$$(\triangleright\nu) \qquad \frac{M \triangleright M'}{M(N) \triangleright M'(N)}$$

$$(\triangleright\xi) \qquad \frac{M \triangleright M'}{(\lambda x.M) \triangleright (\lambda x.M')}.$$

Note that these are the same as the axioms and rules of inference of the λ-calculus, except that '$=$' has been replaced by '$\triangleright$' and the rule (σ) has been omitted. The appropriate definitions of *derivation* and *theory* for λ-reduction should be obvious. We will use '$\vdash$' to denote deducibility in either theory, since the sets of formulas for the two theories are disjoint. Now, what we want to prove is

Theorem 1 (Church-Rosser Theorem for λ-calculus). *If $\vdash M = N$, then there is a term T such that $\vdash M \triangleright T$ and $\vdash N \triangleright T$.*

As we shall see at the end of this section, it is a straightforward matter to get a consistency proof for the λ-calculus from this theorem, and this is of course the reason anyone bothered to prove it in the first place. But as was mentioned above, our interest is more in the techniques and concepts used in this proof.

The bulk of our efforts will be devoted to proving the

Lemma 1 (Church-Rosser Theorem for λ-reduction). *The relation $\triangleright$ has the* **Church-Rosser property**, *i. e. if $\vdash P \triangleright M$ and $\vdash P \triangleright N$, then there exists a term T such that $\vdash M \triangleright T$ and $\vdash N \triangleright T$.*

Proof of the theorem (assuming the lemma). Suppose we have Lemma 1. Then we can inductively define an equivalence relation, which we denote by '$\approx$', to be the least set such that

(1) $\vdash M \triangleright N \implies M \approx N$

(2) $M \approx N \implies N \approx M$

(3) $M \approx N, N \approx L \implies M \approx L$.

Now suppose $M \approx N$. We show inductively that there is a term T such that $\vdash M \triangleright T$ and $\vdash N \triangleright T$. If $M \triangleright N$, take $T \equiv N$. If $N \approx M$, we get T by the inductive hypothesis. So suppose $M \approx L$ and $L \approx N$. By the inductive hypothesis we have two terms T_1 and T_2 such that

$$\vdash M \triangleright T_1 \text{ and } \vdash L \triangleright T_1, \qquad \text{and} \qquad \vdash N \triangleright T_2 \text{ and } \vdash L \triangleright T_2.$$

But $\vdash L \triangleright T_1$ and $\vdash L \triangleright T_2$ give us a T such that $\vdash T_1 \triangleright T$ and $\vdash T_2 \triangleright T$, by lemma 1. The result follows by $(\triangleright\tau)$. All that is required to complete a proof of the theorem are a couple of simpie inductive arguments which show that $M \approx N \iff \vdash M = N$. $\square$

So we turn now to the task of proving the lemma. The proof we give is a translation of a proof due to Per Martin-Löf, who extended a result originally due to W. Tait. Our translation allows us to work with theories rather than with reductions, which is more convenient for the use we will make in later parts of this thesis of the concepts involved in this proof. We employ a strategy similar to that used in appendix 1 of [Hindley and Seldin], though we also borrow liberally from chapter three of [Barendregt 1981].

Proof of the lemma. It will useful to begin with some terminology. A λ-term of form $(\lambda x.M)(N)$ is called a *redex*. If $(\lambda x.M)(N)$ is a redex, then $M[x/N]$ is called its *contractum*. Note that each redex–contractum pair gives rise to an instance of axiom $(\triangleright\beta)$. We need to define a more restricted relation than $\triangleright$.

Definition 8. Let '$\ggg$' denote the smallest relation such that

(i) every redex–contractum pair is included in it

(ii) $$M \ggg N \implies Z(M) \ggg Z(N)$$

(iii) $$M \ggg N \implies M(Z) \ggg N(Z)$$

(iv) $$M \ggg N \implies (\lambda x.M) \ggg (\lambda x.N).$$

This relation is what [Barendregt 1981] calls the *compatible closure* of the redex–contractum relation. Obviously, $M \ggg N \implies \vdash M \triangleright N$. Suppose P is a term in which $(\lambda x.M)(N)$ occurs as a subterm. Use '$P[((\lambda x.M)(N))/(M[x/N])]$' to denote a term which results when we replace an occurrence of $(\lambda x.M)(N)$ by $M[x/N]$. We call $P[((\lambda x.M)(N))/(M[x/N])]$ a *contraction of* $(\lambda x.M)(N)$ *in* P. It is easy to see that $P \ggg P[((\lambda x.M)(N))/(M[x/N])]$.

Definition 9. Let R and S be occurrences of redexes in P, and let P' be the contraction of R in P. The *residuals of* S *with respect to* R are the following redexes in P':

(1) If R and S do not overlap in P, then contracting leaves S unchanged, so S is the residual in P'.

(2) If $R \equiv S$, then S has no residual in P' (since it is contracted out of existence).

(3) If R is a subterm of S and $R \not\equiv S$, then S has form $(\lambda x.M)(N)$ and R is in M or in N. Contracting R changes M to M' or it changes N to N' the residual of S in P' is either $(\lambda x.M')(N)$ or $(\lambda x.M)(N')$.

(4) If S is a subterm of R and $S \not\equiv R$, then R has the form $(\lambda x.M)(N)$ and

S is in M or in N. Contracting R changes $(\lambda x.M)(N)$ to $M[x/N]$. If S is in M, then the residual of S is either S or it is $S[x/N]$ (bear in mind the variable conventions here). If S is in N, there is an occurrence of S in each substituted N, and these are the residuals of S.

Note that only the second half of (4) allows for the possibility that S has more than one residual.

What we are trying to prove here is that the class of proofs in the theory of λ-reduction has a particular property (i. e. the Church-Rosser property). Our strategy will be to isolate a subclass of the proofs for which it is straightforward to show that it is Church-Rosser, then we will reduce the general case to this subclass. Note first of all that for any contraction of a redex R in P there is a proof which establishes $\vdash P \triangleright P'$. Indeed, we introduced the relation $\ggg$ precisely so that this would be obvious. This will be the building block for the type of proof we will focus on.

Definition 10.

(1) Let $R_1, \ldots, R_n$ be occurrences of redexes in a term P. R_i is *minimal* if and only if it properly contains no other R_j.

(2) We say that a term Q is *obtainable from P by a minimal complete derivation (MCD) with respect to* $\{R_1, \ldots, R_n\}$ if Q can be obtained from P by the following process. First contract a minimal R_i (say, for convenience, $i = 1$). Then there are at most $n-1$ residuals $R'_2, \ldots, R'_n$. Contract any minimal R'_j. This leaves at most $n-2$ residuals. Repeat until there are no residuals left. Each of these contractions gives rise to a proof. We get an MCD with

respect to this set of redexes by concatenating these proofs.

(3) We write '$P \ggg_{MCD} Q$', and say that Q is *minimal-completely derivable from* P, if there exists a set $\{R_1, \ldots, R_n\}$ of redexes which occur in P such that Q is obtainable by an MCD with respect to $\{R_1, \ldots, R_n\}$.

Comments.

(1) A single contraction is an (MCD) with respect to a one member set of redexes.

(2) The relativity to the set of redexes is clearly important, since in a term of form $(\lambda x.M)((\lambda y.N)(P))$ we obviously get different results from the singleton sets containing either of the two redexes which occur in that term. The relation $\ggg_{MCD}$ is also not transitive. For example, starting from the term $(\lambda x.x(y))((\lambda z.z))$ we can have only one MCD, namely that with respect to $\{(\lambda x.x(y))((\lambda z.z))\}$ from which we get the term $(\lambda z.z)(y)$. From *that* term we get with respect to $\{(\lambda z.z)(y)\}$ the term y, but there is no MCD that gets us this term from $(\lambda x.x(y))((\lambda z.z))$.

(3) Every non-empty set of redexes has a minimal member.

(4) It is straightforward to show that

$$M \ggg_{MCD} M', N \ggg_{MCD} N' \implies M(N) \ggg_{MCD} M'(N').$$

(5) Keeping in mind the variable conventions, it is not hard to see that if Q and Q' can both be obtained from P by MCD's with respect to the same set of redexes $\{R_1, \ldots, R_n\}$ which occur in P, then $Q \equiv Q'$.

The next step is to prove the necessary but not especially interesting

Sublemma 1. $M \ggg_{MCD} M', N \ggg_{MCD} N' \implies M[x/N] \ggg_{MCD} M'[x/N']$.

Proof of sublemma 1. We argue inductively on the number of atoms in M. Let $R_1, \ldots, R_n$ be the redexes contracted in the MCD of M' from M.

(1) $M \equiv x$. Then $n = 0$ and $M' \equiv x$, so $M[x/N] \equiv N \ggg_{MCD} N' \equiv M'[x/N']$.

(2) $x \notin \mathbf{FV}(M)$. It follows easily from the definition of substitution that $x \notin \mathbf{FV}(M')$, so

$$M[x/N] \equiv M \ggg_{MCD} M' \equiv M'[x/N'].$$

(3) $M \equiv (\lambda y.M_1)$. Then $R_1, \ldots, R_n$ are subterms of M_1, so M' has form $(\lambda y.M_1')$ where $M_1 \ggg_{MCD} M_1'$. Hence

$$\begin{aligned} M[x/N] &\equiv (\lambda y.M_1)[x/N] \\ &\equiv (\lambda y.M_1[x/N]) && \text{since } y \notin \mathbf{FV}(N) \cup \{x\} \\ &\ggg_{MCD} (\lambda y.M_1'[x/N']) && \text{by inductive hypothesis} \\ &\equiv M'[x/N'] && \text{since } y \notin \mathbf{FV}(N') \cup \{x\}. \end{aligned}$$

(4) $M \equiv M_1(M_2)$ and each R_i is in M_1 or in M_2. Then $M' \equiv M_1'(M_2')$ where $M_1 \ggg_{MCD} M_1'$ and $M_2 \ggg_{MCD} M_2'$. So

$$\begin{aligned} M[x/N] &\equiv M_1[x/N](M_2[x/N]) \\ &\ggg_{MCD} M_1'[x/N'](M_2'[x/N']) && \text{by inductive hypothesis and com-} \\ & && \text{mer . (4) above} \\ &\equiv M'[x/N']. \end{aligned}$$

(5) $M \equiv (\lambda y.L)(Q)$ and one R_i (say R_1) is M itself. The residual of R_1 will then be contracted last, so, where $L \ggg_{MCD} L'$ and $Q \ggg_{MCD} Q'$, the

MCD of M' from M has the form

$$M \equiv (\lambda y.L)(Q) \ggg_{MCD} (\lambda y.L')(Q')$$
$$\ggg L'[y/Q'] \equiv M'.$$

By the inductive hypothesis we know that $L[x/N] \ggg_{MCD} L'[x/N']$ and $Q[x/N] \ggg_{MCD} Q'[x/N']$. So

$$\begin{aligned} M[x/N] &\equiv (\lambda y.L[x/N])(Q[x/N]) && \text{since } y \notin \mathbf{FV}(N) \cup \{x\} \\ &\ggg_{MCD} (\lambda y.L'[x/N'])(Q[x/N']) \\ &\ggg L'[x/N'][y/Q'[x/N']] \\ &\equiv L'[y/Q'][x/N'] \equiv M'[x/N'], \end{aligned}$$

relying on a couple of simple corollaries to the definition of substitution and on the variable conventions. But this is clearly an MCD. □

Sublemma 1 is necessary for us to prove cases four and five of the rather more interesting

Sublemma 2. *The relation* $\ggg_{MCD}$ *has the Church-Rosser property, i.e. if* $P \ggg_{MCD} A$*, and* $P \ggg_{MCD} B$*, then there is a term* T *such that* $A \ggg_{MCD} T$ *and* $B \ggg_{MCD} T$.

Proof of Sublemma 2. By induction on the complexity of P.

(1) $P \equiv x$. Then $A \equiv B \equiv P$, so put $T \equiv P$.

(2) $P \equiv (\lambda x.P_1)$. Then all redexes in P are in P_1, so $A \equiv (\lambda x.A_1)$ and $B \equiv (\lambda x.B_1)$, where $P_1 \ggg_{MCD} A_1$ and $P_1 \ggg_{MCD} B_1$. By the inductive

hypothesis there is a T_1 such that $A_1 \ggg_{MCD} T_1$ and $B_1 \ggg_{MCD} T_1$. Put $T \equiv (\lambda x.T_1)$.

(3) $P \equiv P_1(P_2)$ and all the redexes developed in the MCDs are in P_1 or P_2. Then the inductive hypothesis gives us T_1 and T_2 such that we can take $T \equiv T_1(T_2)$.

(4) $P \equiv (\lambda x.M)(N)$ and only one of the given MCDs concludes by contracting the residual of P (say $P \ggg_{MCD} A$). Then that MCD has the form

$$
\begin{aligned}
P &\equiv (\lambda x.M)(N) \\
&\ggg_{MCD} (\lambda x.M')(N') \qquad \text{(where } M \ggg_{MCD} M', N \ggg_{MCD} N'\text{)} \\
&\ggg M'[x/N'] \equiv A.
\end{aligned}
$$

The other MCD has the form

$$
\begin{aligned}
P &\equiv (\lambda x.M)(N) \\
&\ggg_{MCD} (\lambda x.M'')(N'') \qquad \text{(where } M \ggg_{MCD} M'', N \ggg_{MCD} N''\text{)} \\
&\equiv B.
\end{aligned}
$$

The induction hypothesis gives us M^+ and N^+ such that

$$
M' \ggg_{MCD} M^+ \qquad M'' \ggg_{MCD} M^+
$$

$$
N' \ggg_{MCD} N^+ \qquad N'' \ggg_{MCD} N^+.
$$

Put $T \equiv M^+[x/N^+]$. Then $A \equiv M'[x/N'] \ggg_{MCD} M^+[x/N^+]$, by Sub-lemma 1. On the other hand,

$$
\begin{aligned}
B &\equiv (\lambda x.M'')(N'') \\
&\ggg_{MCD} (\lambda x.M^+)(N^+) \qquad \text{by comment (4)} \\
&\ggg M^+[x/N^+],
\end{aligned}
$$

which gives us an MCD as required.

(5) $P \equiv (\lambda x.M)(N)$ and both the given MCDs contract the residual of P.

Then these MCDs have the forms

$$\begin{array}{ll} P \equiv (\lambda x.M)(N) & P \equiv (\lambda x.M)(N) \\ \ggg_{MCD} (\lambda x.M')(N') & \ggg_{MCD} (\lambda x.M'')(N'') \\ \ggg M'[x/N'] \equiv A. & \ggg M''[x/N''] \equiv B. \end{array}$$

The induction hypothesis gives us M^+ and N^+ such that

$$\begin{array}{ll} M' \ggg_{MCD} M^+ & M'' \ggg_{MCD} M^+ \\ N' \ggg_{MCD} N^+ & N'' \ggg_{MCD} N^+. \end{array}$$

Put $T \equiv M^+[x/N^+]$. Then by Sublemma 1 we can argue as at the end of case (4) above to show that we have MCDs $A \equiv M'[x/N'] \ggg_{MCD} M^+[x/N^+]$ and $B \equiv M''[x/N''] \ggg_{MCD} M^+[x/N^+]$ as required. □

Now there are only two steps remaining to prove the lemma. We need

Sublemma 3. *If a binary relation has the Church-Rosser property, then the transitive closure of that relation has the Church-Rosser property.* □

The proof is a trivial induction which is left for the reader. We then show

Sublemma 4. $\rhd = \{ \langle P, Q \rangle \mid\Vdash P \rhd Q \}$ *is the transitive closure of the relation* $\ggg_{MCD}$.

Proof of Sublemma 4. We have already noted that $\ggg_{MCD} \subseteq \rhd$. Now, if we write '$\ggg_=$' for the reflexive closure of $\ggg$, then $\ggg_= \subseteq \ggg_{MCD}$. This is clear since the $M \ggg_= M$ and $(\lambda x.M)(N) \ggg_= M[x/N]$ cases are the zero and one step MCDs, and the closure of $\ggg_{MCD}$ under the rules (ii), (iii) and, (iv) of the definition of $\ggg$ is obvious.

So we have $\ggg_= \subseteq \ggg_{MCD} \subseteq \rhd$. But $\ggg_=$ is just the relation $\rhd$ with the rule $(\rhd_T)$ omitted. So $\rhd$ is the transitive closure of $\ggg_=$. But then $\rhd$ is also the transitive closure of $\ggg_{MCD}$, and we are done. □

We have thus shown that the relation $\rhd$ is has the Church-Rosser property, which completes the proof of Lemma 1, and so also of the theorem. □

3.2 Consistency of the λ-Calculus.

As was mentioned above, the Church-Rosser Theorem almost immediately yields a consistency proof for the λ-calculus. Since we have no notion of negation defined in the untyped λ–calculus, and as we shall see below we cannot have one, we will say that a λ-theory is *inconsistent* if every closed formula is provable in that theory and that it is *consistent* otherwise. We also need the notion of a normal form for a λ–term. A λ–term is *in normal form* if it has no subterm that is a redex. A λ–term *has a normal form* if there is a λ-term N such that N is in normal form and $\vdash M = N$.

We can now use the Church-Rosser Theorem to get the following

Corollaries.

(1) *If a λ-term P has two normal forms A and B, then $A \equiv B$.*

(2) *Let M, N be distinct λ-terms in normal form. Then $\nvdash M = N$.*

(3) *Pure λ-calculus is consistent.*

Proofs.

(1) Recall that $\vdash M = N \iff M \approx N$. By the theorem we have a term T such that $\vdash A \triangleright T$ and $\vdash B \triangleright T$. Since neither A nor B contains a redex, $A \equiv T \equiv B$.

(2) If $\vdash M = N$, then M has two normal forms, namely M and N, so $M \equiv N$.

(3) Define the two combinators

$$\mathbf{K} \equiv (\lambda x.(\lambda y.x)), \qquad \mathbf{S} \equiv (\lambda x.[\lambda y.(\lambda z.x(z)[y(z)])]).$$

These are both terms in normal form, and they are obviously non-congruent, so $\nvdash \mathbf{K} = \mathbf{S}$ by (2). □

3.3 The $\lambda\eta$-Calculus.

We take this opportunity to list (without proofs) some facts about an extension of the λ-calculus which will sometimes be mentioned in the discussion below.

Definition 11.

(1) $\lambda + (\text{ext})$ is the theory which results from adding the inference rule

$$(\text{ext}) \qquad \frac{M(x) = N(x)}{M = N} \qquad \text{provided } x \notin \mathbf{FV}(M) \cup \mathbf{FV}(N)$$

to rules of inference of the λ–calculus.

(2) $\lambda\eta$ is the theory which results by adding the axiom scheme

$$(\eta\text{-conversion}) \qquad (\lambda x.M)(x) = M \qquad \text{provided } x \notin \mathbf{FV}(M)$$

to the axioms of the λ-calculus.

Theorem 2. *The theories $\lambda + (ext)$ and $\lambda\eta$ are equivalent.* $\square$

For a proof, see [Barendregt 1981, p.32] or [Hindley and Seldin, p.74]. We will discuss $\lambda\eta$, as is usual in the literature. We will need to adapt some of our earlier terminology. An *η-redex* is a redex or is a λ–term of form $(\lambda x.M)(x)$ with $x \notin \mathbf{FV}(M)$. If $(\lambda x.M)(x)$ is an η-redex, then M is its *η-contractum*. If T is an η-redex not of this form, then its *η-contractum* is its contractum.

The *theory of η-reduction* is the theory which results when we add the axiom scheme

$$(\lambda x.M)(x) \rhd M \qquad \text{provided } x \notin \mathbf{FV}(M)$$

to the theory of λ–reduction. We write '$\eta \vdash M = N$' and '$\eta \vdash M \rhd N$' for the obvious claims about provability in $\lambda\eta$ and in the theory of η-reduction.

Lemma 2. *The theory of η–reduction is Church-Rosser, i.e. if $\eta \vdash P \rhd M$ and $\eta \vdash P \rhd N$, then there is a term T such that $\eta \vdash M \rhd T$ and $\eta \vdash N \rhd T$.* $\square$

Theorem 3 (Church-Rosser theorem for $\lambda\eta$). *If $\eta \vdash M = N$, then there is a term T such that $\eta \vdash M \rhd T$ and $\eta \vdash N \rhd T$.* $\square$

Industrious readers can reconstruct proofs of these two claims along the lines of the proofs above without difficulty from [Barendregt 1981 §3.3]. We say that a λ-term is *in η-normal form* if it has no subterm which is an η-redex. A λ-term M *has an η-normal form* if there is a term N such that N is in η-normal form and $\eta \vdash M = N$.

Corollaries.

(1) *Let M and N be distinct terms in η-normal form. Then* $\nvdash M = N$.

(2) $\lambda\eta$ *is consistent.* □

We can use the same combinators **K** and **S** we used above to prove consistency of the pure λ–calculus to prove consistency here. One characteristic which $\lambda\eta$ has which the λ-calculus does not share is given by the following theorem, which is theorem 2.1.40 of [Barendregt 1981].

Theorem 4. $\lambda\eta$ *is Hilbert-Post complete, i.e. if M and N each have an η-normal form, then either* $\eta \vdash M = N$, *or* $\lambda\eta \cup \{M = N\}$ *is inconsistent.* □

3.4 Fixpoint Theorems and Paradoxes. Another of the classic theorems of the λ-calculus will allow us to demonstrate why some limits apply when we try to combine propositional logic and λ–calculus.

Theorem 5 (Fixpoint Theorem).

(1) *For every λ-term M there is a term P such that* $\vdash M(P) = P$.

(2) *Furthermore, there is a combinator* **Y** *which finds these fixed points, that is, a Y such that for every term M,* $\vdash M(Y(M)) = Y(M)$.

(3) *Indeed, we can strengthen this to the claim that* $\vdash M(Y(M)) \triangleright Y(M)$.

Proof. All three parts follow from a demonstration that the combinator (due to Turing)

$$[\lambda z.[\lambda x.x(z(z)(x))]]([\lambda z.[\lambda x.x(z(z)(x))]])$$

has the last property. See [Hindley and Seldin p. 33].

We satisfy ourselves with showing that the more familiar fixpoint combinator (due to Curry)

$$[\lambda x.[\lambda y.x(y(y))]([\lambda y.x(y(y))])]$$

gives us the first two parts of the theorem. So write 'W' for '$[\lambda y.x(y(y))]$'. Then

$$\begin{aligned}
(\lambda x.W(W))(M) &= W(W)[x/M] \\
&\equiv (\lambda y.M(y(y)))((\lambda y.(M(y(y)))) \\
&= M(y(y))[y/(\lambda y.M(y(y)))] && \text{by } (\beta) \\
&\equiv M((\lambda y.M(y(y)))((\lambda y.M(y(y)))) \\
&\equiv M((\lambda y.x(y(y)))((\lambda y.x(y(y))))[x/M]) \\
&= M((\lambda x.W(W))(M)) && \text{by } (\beta). \quad \Box
\end{aligned}$$

Now, the most obvious way to attempt to combine the λ-calculus with a system of logic is to introduce constants for the propositional connectives and quantifiers, and to take the postulates necessary for first order logic—either intuitionistic or classical. However, it is not hard to see that we are quickly going to have problems with this approach, before we even get to the stage of trying to introduce quantifiers. For example, if we introduce **not** as a constant, the fixpoint theorem tells us that there is a λ-term P such that $\vdash P = \mathbf{not}P$. Clearly nothing deserving of the name 'negation' can satisfy this property.

It is also worth noting that we can't simply throw out negation and try to satisfy ourselves with an implicational logic, either. For if we want to introduce a constant

'⊃' for implication, we will certainly require, either as axioms or theorems, all formulas of the form

$$(*) \qquad (X \supset (X \supset Y)) \supset (X \supset Y)$$

and the rule

$$(\mathrm{MP}) \qquad \frac{X \supset Y \qquad X}{Y}.$$

But if we have these and the seemingly indispensable rule

$$(\mathrm{eq}) \qquad \frac{X \qquad X = Y}{Y},$$

we also have $\vdash Z$ for every Z. To see this, let $\mathbf{Y}$ be a fixpoint operator and define

$$X \equiv \mathbf{Y}(\lambda z.z \supset (z \supset Z)),$$

where $z \notin \mathbf{FV}(Z)$. Since $\mathbf{Y}$ is a fixpoint operator, $(\lambda z.z \supset (z \supset Z))(X) = X$. This gives us

$$X = X \supset (X \supset Z) = (X \supset (X \supset Z)) \supset (X \supset Z).$$

So,

$(X \supset (X \supset Z)) \supset (X \supset Z)$	by $*$
$X \supset (X \supset Z)$	by (eq)
X	by (eq)
$X \supset Z$	by MP
Z	by MP.

This result is called "Curry's Paradox" (and the fixpoint combinator used in the proof or Theorem 5 is called by Curry the "Paradoxical Combinator"). The usual approach of those, like Curry, who nonetheless want to use λ-calculus as a basis for logic is to enforce a distinction between those terms which can represent propositions and those which cannot, which is to say that in effect they give up the attempt to use the untyped λ–calculus. The result of their efforts is known as "illative combinatory logic". Since our focus is the untyped calculus, we will not pursue illative combinatory logic here. In particular, the model theory we will discuss in the next chapter is for the λ-calculus without this type distinction.

Of course, this does not entail that we cannot use the usual equipment of logic in our metalanguage to say things like

$$M = N \implies M(N) = N(M)$$

for

$$\vdash M = N \quad \text{implies} \quad \vdash M(N) = N(M).$$

Indeed, this is the common practice in the literature. Clearly, we could formalize this practice—we could treat the λ–equations as atomic formulas and use the usual definitions from logic to generate the formulas out of these atoms. This is just what [Hindley and Seldin] do. The contradictions result when we try to treat the logical operators as λ–terms. As long as we enforce a strict separation between the connectives and the λ-equations—in particular. so long as we ensure that the abstraction operator binds only within the λ-terms and not across logical connectives—there is no problem. It is this need for a strict separation which makes the abstraction operator λ a non-candidate for the status of *term forming operator* as we will use that phrase in later chapters. For the operators we will be concerned with must be such that we can use them to form terms out of arbitrary formulas—*a fortiori* out of formulas which contain connectives.

On the other hand, the type-free λ-calculus is not entirely devoid of features corresponding to those found in predicate logic. Consider, for example, the question of the quantifiers. It might well seem that one could define

$$\forall x(M = N) \iff (\lambda x.M = \lambda x.N).$$

As we will see in the next chapter, this is not something we can do in general, because the $\Longrightarrow$ direction is stronger than the (ξ)-rule. However, if we are willing to make a simplifying assumption, which we will call *weak extensionality*, we can make this definition and so the abstraction operator has the internal strength of a universal quantifier.

CHAPTER IV

λ—CALCULUS: SEMANTICS

The first models of the λ–calculus to be developed were the so-called term models. These models, which are syntactical in nature, were proven to be non-trivial by the Church–Rosser Theorem (theorem 3.7), which was first published in 1936. Indeed, the untyped λ–calculus as a formal system had been fairly thoroughly investigated by 1940. However, it was not until about 1970 that mathematical (i. e. non-syntactical) models began to be developed. Since then, though, the development of this model theory has been rapid.

The main barrier to the development of these mathematical models stems from the fact that in the untyped λ–calculus if f and g are any terms, then $f(g)$ (f applied to g) is always defined. The λ-terms are intended to be regarded as algorithms or rules, which can themselves be considered as data, and so given as input to other algorithms. Now if we try to develop an interpretation of the λ–calculus by making the obvious move of interpreting algorithms by set theoretic functions, we quickly run into trouble. Since every term can occur as both function and input to a function, we will ideally want to have, where $\mathcal{A}$ is the domain of our interpretation and $[\mathcal{A} \to \mathcal{A}]$ is the space of functions from $\mathcal{A}$ to $\mathcal{A}$, $[\mathcal{A} \to \mathcal{A}] \cong \mathcal{A}$. Equally clearly, by Cantor's theorem this is something we can't have if we interpret $[\mathcal{A} \to \mathcal{A}]$ as the full set theoretic function space.

The difficulty, then, was to find suitable mathematical objects and maps between them so that for a set of such objects $\mathcal{X}$ the space of such maps $[\mathcal{X} \to \mathcal{X}]$ could be embedded back into $\mathcal{X}$. The first to find such objects and such maps was Dana Scott (see [Scott 1972]), who constructed the object D_∞ as a suitable limit in the category of continuous lattices. Shortly thereafter Plotkin (see [Plotkin 1972]) and Scott independently discovered the graphmodel $\mathcal{P}\omega$ which we will describe at the end of this chapter.

The invention of cartesian closed categories ([Lawvere 1964], [Eilenberg and Kelly]) and the discovery that there is an equivalence of categories between the category of cartesian closed categories and the category of *typed* λ–calculi with surjective pairing[1] (due to Lambek, proved in detail in [Lambek and Scott]) led to another approach to the model theory for the *type–free* λ–calculus. Lambek and Dana Scott discovered that the category of C-monoids (essentially one-object cartesian closed categories without terminal object) is isomorphic with the category of untyped λ–calculi which meet the further condition of having surjective pairing. This yields a handy way of presenting many of the models of the type free calculus.[2]

However, all these mathematical models so far mentioned interpret any λ–abstraction term in a way such that its interpretation is completely determined by the applicative behaviour of the term from which it is abstracted. That is, the

[1] By 'surjective pairing' is meant that we have three combinators satisfying the obvious conditions for ordered pairing and projection functions and which meet the further condition that, for each c, $\langle \pi_1(c), \pi_2(c) \rangle = c$. This last condition, surjectivity, cannot hold in the pure λ–calculus, typed or untyped.

[2] However, we shall not discuss C-monoids below when we give a category theoretic semantics for the λ–calculus. Instead, we will give a semantics for the pure λ–calculus since the Church-Rosser Theorem fails for the λ–calculus with surjective pairing, and it is easier to present the difference between extensional and non-extensional models in the semantics we will use here.

interpretation of $(\lambda x.M)$ depends, where D is the domain of our model, and where $\mathbf{d}$ is a name for d, on the interpretation of $M[x/\mathbf{d}]$ for all $d \in D$ and on nothing else. This has the consequence that, writing $\mathcal{M}$ for a model and $|\mathcal{M}|$ for its domain,

$$(\forall d \in |\mathcal{M}|.\, \mathcal{M} \models M[x/\mathbf{d}] = N[x/\mathbf{d}]) \implies \mathcal{M} \models (\lambda x.M) = (\lambda x.N).$$

A model satisfying this condition is usually called *weakly extensional*.

Weak extensionality does not seem to be an unnatural condition for models of the λ–calculus to meet if we regard it as a calculus of functions. However, many authors, e. g. Barendregt and Koymans, object to this by pointing out that the structure of all closed λ–terms does not satisfy this condition. This, they contend, should count as a model of the λ–calculus, and indeed it is one of the term models we mentioned above as the first models to be developed. Furthermore, if we regard the λ–calculus as an attempt at a calculus of algorithms rather than of functions, as its inventors intended, the requirement seems out of place. If we think on analogy with Turing machines, where each algorithm is a program which has a coding which can be fed as input into other programs, we clearly will not have the condition of weak extensionality met, since the same input-output behaviour can result from algorithms that are syntactically very different, and so end up with different codings.

Such considerations have led researchers to an investigation of whether there are interesting mathematical models of the λ–calculus which are not weakly extensional. As a result, it has become common to make the following terminological stipulations: a λ–*algebra* is a structure such that the equations provable in the λ-calculus are all valid in it, while the name λ–*model* is reserved for the subclass of the λ–algebras which are weakly extensional. (see, e. g. [Koymans p. 3–4]). Finally, to explain why the word 'weakly' keeps popping up, we mention that the *extensional λ–models* are the ones in which all the elements of the domain (i. e. not just the abstraction terms) are determined by their applicative behaviour, that is,

$$\forall a, b \in |\mathcal{M}|, ([\forall d \in |\mathcal{M}|.\, a(d) = b(d)] \implies a = b).$$

Our discussion of the model theory of the λ–calculus will proceed as follows. First, we will give some definitions which mirror fairly closely the rules of the λ-calculus. We will use these to give the background we need to be able to discuss the term models. We will use the term models to generate various completeness results, and to illustrate some points about extensionality which will be used to justify distinguishing between *λ-models* and *λ-algebras*.

The most natural description of the mathematical models mentioned above is a category theoretic one: each reflexive object in a cartesian closed category gives rise to a λ-algebra, and this λ–algebra is a λ–model if the reflexive object has enough points. In order to make the category theoretic description as readable as possible for those not acquainted with category theory, we will work up to it in stages by first giving another mathematical semantics (usually called "first order semantics") for λ–calculus. This semantics employs algebraic concepts which are likely to be more familiar to those whose background in logic includes only the usual introductory courses taught in philosophy departments. First order semantics is fairly easily shown to yield a stock of structures which is categorically isomorphic to the original syntactical account. This will enable us to give the category theoretic description of the models and the algebras by relating it to the first order structures. When we have the tools available we will describe in category theoretic terms the relationship between the algebras and the models.

After all this we will be able to give an efficient and straightforward account of the graph model for the λ-calculus.

1. "Syntactical" λ-Semantics

We need some terminology to begin with. $\mathcal{M} = \langle X, \circ \rangle$ is an *applicative structure* if X is a set and $\circ : X^2 \to X$. We sometimes write '$a \in \mathcal{M}$' for '$a \in X$'. It will be convenient to have names for the members of X, so for each $a \in X$, let $\mathbf{a}$ be a constant. Let '$\Lambda(\mathcal{M})$' denote the set of λ-terms with $\{ \mathbf{a} \mid a \in X \}$ as constants. A *valuation in* $\mathcal{M}$ is a map $\varrho : \mathbf{Var} \to X$. If $\varrho : \mathbf{Var} \to X$ is a valuation in $\mathcal{M}$, then for $a \in X$ $\varrho(x/a)$ is the valuation $\varrho' : \mathbf{Var} \to X$ such that $\varrho'(x) = a$ and $\varrho'(y) = \varrho(y)$ for $y \neq x$. We will continue to use 'A','B',..., 'M,' 'N', 'P', ... to denote arbitrary terms and 'x', 'y',... to denote arbitrary variables in $\Lambda(\mathcal{M})$.

Definition 1. Let $\mathcal{M} = \langle X, \circ \rangle$ be an applicative structure. An *interpretation* for $\mathcal{M}$ is a map $\| \cdot \| : \Lambda(\mathcal{M}) \times \{ \varrho \mid \varrho : \mathbf{Var} \to X \} \to X$ such that for any $\varrho : \mathbf{Var} \to X$,

(1) $\|x\|^{\varrho}_{\mathcal{M}} = \varrho(x)$ for $x \in \mathbf{Var}$

(2) $\|\mathbf{a}\|^{\varrho}_{\mathcal{M}} = a$ for $a \in X$

(3) $\|P(Q)\|^{\varrho}_{\mathcal{M}} = \|P\|^{\varrho}_{\mathcal{M}} \circ \|Q\|^{\varrho}_{\mathcal{M}}$

(4) $\|(\lambda x.P)\|^{\varrho}_{\mathcal{M}} \circ a = \|P\|^{\varrho(x/a)}_{\mathcal{M}}$

(5) $\varrho \restriction \mathbf{FV}(M) = \varrho' \restriction \mathbf{FV}(M) \implies \|M\|^{\varrho}_{\mathcal{M}} = \|M\|^{\varrho'}_{\mathcal{M}}$.

$\mathcal{M} = \langle X, \circ, \|\cdot\|\rangle$ is a *syntactical applicative structure* if $\langle X, \circ\rangle$ is an applicative structure and $\|\cdot\|$ is an interpretation for it.

It is not difficult to see where these various clauses come from. Obviously we will want an interpretation to interpret every term, and (1) is the obvious definition for the case $M \equiv x$, while (3) is the obvious definition for $M \equiv P(Q)$. As for (2), it too is standard if we want the convenience of having names for the elements of the structure. No one will be surprised by (5). Finally, (4) is included to ensure that the structure interprets the λ-abstraction terms in accordance with the intuitive meaning described for them in chapter 3.

We will say that a syntactical applicative structure $\mathcal{M}$ *satisfies the equation* $M = N$ *with* ϱ, and will write '$\mathcal{M}, \varrho \models M = N$' if $\|M\|^{\varrho}_{\mathcal{M}} = \|N\|^{\varrho}_{\mathcal{M}}$. We will write '$\mathcal{M} \models M = N$', and will say $\mathcal{M}$ *satisfies* $M = N$ if $\mathcal{M}, \varrho \models M = N$ for every valuation $\varrho : \mathbf{Var} \to X$ in $\mathcal{M}$. We will use the same names for axioms and rules of inference in the λ-calculus as were used in Chapter 3.

Lemma 1. *Any syntactical applicative structure $\mathcal{M}$ satisfies all instances of the (ρ)-schema. Moreover the class of equations satisfied by $\mathcal{M}$ is closed under the rules (σ), (τ), (μ), and (ν).*

Proof. The instances of the (ρ)-schema and closure under (τ) and (σ) follow from the fact that identity in $\mathcal{M}$ is an equivalence.

If $\mathcal{M} \models M = M'$, then for any $\varrho : \mathbf{Var} \to X$, $\|M\|^{\varrho}_{\mathcal{M}} = \|M'\|^{\varrho}_{\mathcal{M}}$. So taking ($\mu$)

and (ν) in turn, we have

$$\|N(M)\|_{\mathcal{M}}^{\varrho} = \|N\|_{\mathcal{M}}^{\varrho} \circ \|M\|_{\mathcal{M}}^{\varrho} = \|N\|_{\mathcal{M}}^{\varrho} \circ \|M'\|_{\mathcal{M}}^{\varrho} = \|N(M')\|_{\mathcal{M}}^{\varrho} .$$

$$\|M(N)\|_{\mathcal{M}}^{\varrho} = \|M\|_{\mathcal{M}}^{\varrho} \circ \|N\|_{\mathcal{M}}^{\varrho} = \|M'\|_{\mathcal{M}}^{\varrho} \circ \|N\|_{\mathcal{M}}^{\varrho} = \|M'(N)\|_{\mathcal{M}}^{\varrho} . \quad \square$$

All that remains to do to get a soundness proof is to show that the (β)-axioms and the rule (ξ) are valid. Unfortunately, even with (4) of the definition of an interpretation we cannot ensure that the (β)-axioms are satisfied. The problem is that we cannot show, on the basis of the definitions we have so far, that

$$\|(\lambda y.M)[x/N]\|_{\mathcal{M}}^{\varrho} = \|(\lambda y.M)\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})} .$$

The definition we are about to give of a syntactical λ-model will kill two birds with one stone, since it will imply both closure under the (ξ)-rule and validity of the above formula, which will in turn allow us to prove soundness. The definition of a syntactical λ-algebra merely stipulates that (ξ) and (β) must hold.

Definition 2. Let $\mathcal{M}$ be a syntactical applicative structure. $\mathcal{M}$ is a *syntactical λ-algebra* if and only if

$$\vdash M = N \implies \mathcal{M} \models M = N.$$

$\mathcal{M}$ is a *syntactical λ-model* if and only if

$$\text{(w.e.)} \quad \|M\|_{\mathcal{M}}^{\varrho(x/a)} = \|N\|_{\mathcal{M}}^{\varrho(x/a)} \quad \text{for all } a \in X \implies \|(\lambda x.M)\|_{\mathcal{M}}^{\varrho} = \|(\lambda x.N)\|_{\mathcal{M}}^{\varrho} .$$

The Soundness Theorem is thus true by definition for syntactical λ-algebras. To prove it for the syntactical λ-models, we need

Lemma 2. *Let $\mathcal{M}$ be a syntactical λ-model. Then for any valuation $\varrho : \mathbf{Var} \to X$,*

$$\|M[x/N]\|_{\mathcal{M}}^{\varrho} = \|M\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})}.$$

Proof. The proof is an induction on the complexity of M, the only difficult case of which is where $M \equiv (\lambda y.P)$. By the variable conventions $y \not\equiv x$ and $y \notin \mathbf{FV}(N)$. Note that $(\lambda y.P)[x/N] \equiv (\lambda y.P[x/N])$. We must show

$$\|(\lambda y.P[x/N])\|_{\mathcal{M}}^{\varrho} = \|(\lambda y.P)\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})}$$

and our inductive hypothesis is that for each $a \in X$

$$\|P[x/N]\|_{\mathcal{M}}^{\varrho(y/a)} = \|P\|_{\mathcal{M}}^{\varrho(y/a)(x/\|N\|_{\mathcal{M}}^{\varrho})}.$$

Suppose $x \notin \mathbf{FV}(N)$. Since $y \notin \mathbf{FV}(N)$, we have

$$\|P[x/N]\|_{\mathcal{M}}^{\varrho(y/a)(x/\|N\|_{\mathcal{M}}^{\varrho})} = \|P[x/N]\|_{\mathcal{M}}^{\varrho(y/a)} \qquad \text{so}$$

$$\|P[x/N]\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})(y/a)} = \|P\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})(y/a)} \qquad \text{so}$$

$$\|(\lambda y.P[x/N])\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})} = \|(\lambda y.P)\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})} \qquad \text{by (w.e.), so}$$

$$\|(\lambda y.P[x/N])\|_{\mathcal{M}}^{\varrho} = \|(\lambda y.P)\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})}.$$

So suppose $x \in \mathbf{FV}(N)$. Let z be a fresh variable. Then

$$\|(\lambda y.P[x/N])\|_{\mathcal{M}}^{\varrho} = \|(\lambda y.P[x/z][z/N])\|_{\mathcal{M}}^{\varrho}$$
$$= \|(\lambda y.P[x/z])\|_{\mathcal{M}}^{\varrho(z/\|N\|_{\mathcal{M}}^{\varrho})}$$

by the above, since $z \notin \mathbf{FV}(N)$. But since x is not free in z, the lemma holds and we have

$$
\begin{aligned}
\|(\lambda y.P[x/y])\|_{\mathcal{M}}^{\varrho(z/\|N\|_{\mathcal{M}}^{\varrho})} &= \|(\lambda y.P)\|_{\mathcal{M}}^{\varrho(z/\|N\|_{\mathcal{M}}^{\varrho})(x/\|z\|_{\mathcal{M}}^{\varrho(z/\|N\|_{\mathcal{M}}^{\varrho})})} \\
&= \|(\lambda y.P)\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})},
\end{aligned}
$$

since z does not occur in P. □

Theorem 1 (Soundness Theorem for λ-Models). *Let $\mathcal{M}$ be a syntactical λ-model. Then*

$$\vdash M = N \implies \mathcal{M} \models M = N,$$

i.e. $\mathcal{M}$ is a syntactical λ-algebra.

Proof. By Lemma 1, it suffices to show that the (β)-axioms and the (ξ)-rule are valid. For (β) we have

$\|(\lambda x.M)(N)\|_{\mathcal{M}}^{\varrho} = \|(\lambda x.M)\|_{\mathcal{M}}^{\varrho} \circ \|N\|_{\mathcal{M}}^{\varrho}$	by (3) of the definition of an interpretation
$= \|M\|_{\mathcal{M}}^{\varrho(x/\|N\|_{\mathcal{M}}^{\varrho})}$	by (4) of the definition of an interpretation
$= \|M[x/N]\|_{\mathcal{M}}^{\varrho}$	by lemma 2.

For (ξ), suppose $\|M\|_{\mathcal{M}}^{\varrho} = \|N\|_{\mathcal{M}}^{\varrho}$ for all valuations ϱ. Then $\|M\|_{\mathcal{M}}^{\varrho(x/a)} = \|N\|_{\mathcal{M}}^{\varrho(x/a)}$, since, for all $a \in X$, $\varrho(x/a)$ is a valuation in $\mathcal{M}$. So $\|(\lambda x.M)\|_{\mathcal{M}}^{\varrho} = \|(\lambda x.N)\|_{\mathcal{M}}^{\varrho}$ by (w.e.). □

In order to get a completeness proof we will now consider term models, which will serve to give canonical interpretations in a sense that will become clear below.

Let $\mathbf{T}$ be a $\Lambda(\mathbf{C})$-theory. We define the relation $=_{\mathbf{T}}$ on $\Lambda(\mathbf{C})$ and the relation $=_{\mathbf{T}^0}$ on $\Lambda^0(\mathbf{C})$ by

$$M =_{\mathbf{T}} N \iff \mathbf{T} \vdash M = N \qquad \text{and} \qquad M =_{\mathbf{T}^0} N \iff \mathbf{T} \vdash M = N.$$

Put $[M]_{\mathbf{T}} = \{ N \in \Lambda(\mathbf{C}) \mid \mathbf{T} \vdash M = N \}$ for all $M \in \Lambda(\mathbf{C})$ and put, for $M \in \Lambda^0(\mathbf{C})$, $[M]_{\mathbf{T}^0} = \{ N \in \Lambda^0(\mathbf{C}) \mid \mathbf{T} \vdash M = N \}$. Let $X = \{ [M]_{\mathbf{T}} \mid N \in \Lambda(\mathbf{C}) \}$ and $X^0 = \{ [M]_{\mathbf{T}^0} \mid N \in \Lambda^0(\mathbf{C}) \}$. We now define $\bullet : X^2 \to X$ and $\bullet^0 : X^{0^2} \to X^0$ by

$$[M]_{\mathbf{T}} \bullet [N]_{\mathbf{T}} = [M(N)]_{\mathbf{T}}$$

and

$$[M]_{\mathbf{T}^0} \bullet^0 [N]_{\mathbf{T}^0} = [M(N)]_{\mathbf{T}^0}.$$

It is easy to check that this is all well defined, so we have two applicative structures $\mathcal{M}(\mathbf{T}) = \langle X, \bullet \rangle$ and $\mathcal{M}^0(\mathbf{T}) = \langle X^0, \bullet^0 \rangle$. Obviously we want to define interpretations for these two structures. The first step is to add a constant for each element of the domain. We write '$[\mathbf{M}]_{\mathbf{T}}$' and '$[\mathbf{M}]_{\mathbf{T}^0}$' for the constants which name the equivalence classes $[M]_{\mathbf{T}} \in X$ and $[M]_{\mathbf{T}^0} \in X^0$ respectively. (So in particular we write '$\varrho(\mathbf{x})$' the constant naming $[\varrho(x)]_{\mathbf{T}}$ or $[\varrho(x)]_{\mathbf{T}^0}$, as the case may be.) Instead of working with $\Lambda(\mathbf{C})$ and $\Lambda^0(\mathbf{C})$, we from here on work with the sets of λ-terms $\Lambda(\mathcal{M}(\mathbf{T}))$ and $\Lambda(\mathcal{M}^0(\mathbf{T}))$, which we get by adding the new constants to $\mathbf{C}$ in each case.

We begin by interpreting the closed terms in each of these sets. For $M \in \Lambda^0(\mathcal{M}(\mathbf{T}))$ (respectively, $M \in \Lambda^0(\mathcal{M}^0(\mathbf{T}))$) we denote by '$M'$' the term obtained from M by replacing every constant of the form $[\mathbf{N}]_{\mathbf{T}}$ ($[\mathbf{N}]_{\mathbf{T}^0}$) by N. Here we are relying on the variable conventions to prevent variable collisions. We now put $\|M\| = [M']_{\mathbf{T}}$ ($\|M\| = [M']_{\mathbf{T}^0}$). It is straightforward to confirm that the choice of representative for $[N]_{\mathbf{T}}$ in the last sentence is irrelevant.

We now have an interpretation (in both cases) only for the closed terms. We extend these to interpretations for arbitrary terms as follows: For $M \in \Lambda(\mathcal{M}(\mathbf{T}))$ (resp. $M \in \Lambda(\mathcal{M}^0(\mathbf{T}))$), with $\mathbf{FV}(M) \subseteq \{x_1, \ldots, x_n\}$, and for any $\varrho : \mathbf{Var} \to X$ (resp. $\varrho : \mathbf{Var} \to X^0$), put

$$\|M\|^{\varrho} = \|M[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})]\| \,.$$

Definition 3. $\mathcal{M}(\mathbf{T}) = \langle X, \bullet, \| \cdot \| \rangle$ is *the open term model* of $\mathbf{T}$. $\mathcal{M}^0(\mathbf{T}) = \langle X^0, \bullet^0, \| \cdot \| \rangle$ is *the closed term model* of $\mathbf{T}$.

Lemma 3. *In both the open and closed term models, $\| \cdot \|$ is an interpretation.*

Proof. Conditions (1) and (2) of the definition of an interpretation are true by definition, while (5) is trivial. For (3), let $\varrho : \mathbf{Var} \to X$, let $P, Q \in \Lambda(\mathcal{M}(\mathbf{T}))$, assume $\mathbf{FV}(P) \cup \mathbf{FV}(Q) \subseteq \{x_1, \ldots, x_n\}$, and assume that the constants of the form $[\mathbf{N}]_{\mathbf{T}}$ in P and Q are among $[\mathbf{N_1}]_{\mathbf{T}}, \ldots, [\mathbf{N_m}]_{\mathbf{T}}$. We then have

$$\begin{aligned}
\|P(Q)\|^{\varrho} &= \|P(Q)[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})]\| \\
&= \left[(P(Q)[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})])'\right]_{\mathbf{T}} \\
&= [P(Q)[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n), [\mathbf{N_1}]_{\mathbf{T}}/N_1, \ldots, [\mathbf{N_m}]_{\mathbf{T}}/N_m]]_{\mathbf{T}} \\
&= [P[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n), [\mathbf{N_1}]_{\mathbf{T}}/N_1, \ldots, [\mathbf{N_m}]_{\mathbf{T}}/N_m] \\
&\qquad (Q[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n), [\mathbf{N_1}]_{\mathbf{T}}/N_1, \ldots, [\mathbf{N_m}]_{\mathbf{T}}/N_m])]_{\mathbf{T}}
\end{aligned}$$

by the definition of substitution, so

$$\|P(Q)\|\|^{\varrho} = P[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n), [\mathbf{N_1}]_{\mathbf{T}}/N_1, \ldots, [\mathbf{N_m}]_{\mathbf{T}}/N_m]_{\mathbf{T}}$$
$$\bullet\, Q[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n), [\mathbf{N_1}]_{\mathbf{T}}/N_1, \ldots, [\mathbf{N_m}]_{\mathbf{T}}/N_m]_{\mathbf{T}}$$

$$= P[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})]_{\mathbf{T}} \bullet Q[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})]_{\mathbf{T}}$$
$$= \|P\|^{\varrho} \bullet \|Q\|^{\varrho}.$$

The same proof goes through for the case of $P, Q \in \Lambda(\mathcal{M}^0(\mathbf{T}))$ if we replace the subscripted $\mathbf{T}$'s by $\mathbf{T}_0$'s, and ϱ by a map into X^0.

For (4), let $\varrho : \mathbf{Var} \to X$, let $P \in \Lambda(\mathcal{M}(\mathbf{T}))$, assume $\mathbf{FV}(P) \subseteq \{x, x_1, \ldots, x_n\}$ and assume that the constants of the form $[\mathbf{N}]_{\mathbf{T}}$ in P are among $[\mathbf{N_1}]_{\mathbf{T}}, \ldots, [\mathbf{N_m}]_{\mathbf{T}}$. Note that $\|(\lambda x.P[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n)])\| \in X$. So

$$\|(\lambda x.P)\|^{\varrho} \bullet [Q]_{\mathbf{T}} = \|(\lambda x.P[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n}))]\| \bullet [Q]_{\mathbf{T}}$$
$$= [(\lambda x.P[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n), [\mathbf{N_1}]_{\mathbf{T}}/N_1, \ldots, [\mathbf{N_m}]_{\mathbf{T}}/N_m])]_{\mathbf{T}} \bullet [Q]_{\mathbf{T}}$$
$$= [(\lambda x.P[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n), [\mathbf{N_1}]_{\mathbf{T}}/N_1, \ldots, [\mathbf{N_m}]_{\mathbf{T}}/N_m])(Q)]_{\mathbf{T}}$$
$$= [P[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n), [\mathbf{N_1}]_{\mathbf{T}}/N_1, \ldots, [\mathbf{N_m}]_{\mathbf{T}}/N_m][x/Q]]_{\mathbf{T}}$$
$$= [P[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})][x/Q]]_{\mathbf{T}}$$
$$= \|P\|^{\varrho(x/[Q]_{\mathbf{T}})}.$$

For $P \in \Lambda(\mathcal{M}^0(\mathbf{T}))$ this proof goes through with the same alterations needed in the case of (3). □

Lemma 3 shows that both the open and closed term models are syntactical applicative structures, which makes it easier to prove

Theorem 2. *The open term model is a λ-model.*

Proof. By the above lemmas it suffices to show that (w.e.) is satisfied. So suppose $\|M\|^{\varrho(x/[P]_{\mathbf{T}})} = \|N\|^{\varrho(x/[P]_{\mathbf{T}})}$ for all $[P]_{\mathbf{T}}$. This implies $\|M\|^{\varrho(x/[\mathbf{x}]_{\mathbf{T}})} = \|N\|^{\varrho(x/[\mathbf{x}]_{\mathbf{T}})}$. Assume $(\mathbf{FV}(M) \cup \mathbf{FV}(N)) - \{x\} \subseteq \{x_1, \ldots, x_n\}$, and put $\varrho' = \varrho(x/[\mathbf{x}]_{\mathbf{T}})$. By the definition of $\| \cdot \|^{\varrho}$, this gives us

$$[M[x_1/\varrho'(\mathbf{x_1}), \ldots, x_n/\varrho'(\mathbf{x_n})]]_{\mathbf{T}} = [N[x_1/\varrho'(\mathbf{x_1}), \ldots, x_n/\varrho'(\mathbf{x_n})]]_{\mathbf{T}}\,,$$

so,

$$[M[x_1/\varrho'(x_1), \ldots, x_n/\varrho'(x_n)]]_{\mathbf{T}} = [N[x_1/\varrho'(x_1), \ldots, x_n/\varrho'(x_n)]]_{\mathbf{T}}\,,$$

and so $[M']_{\mathbf{T}} = [N']_{\mathbf{T}}$, where if x is free in M, N, then x is free in M', N'. By the definition of $[\cdot]_{\mathbf{T}}$ and the (ξ)-rule, $[(\lambda x.M')]_{\mathbf{T}} = [(\lambda x.N')]_{\mathbf{T}}$, so

$$[(\lambda x.M[x_1/\varrho'(x_1), \ldots, x_n/\varrho'(x_n)])]_{\mathbf{T}} = [(\lambda x.N[x_1/\varrho'(x_1), \ldots, x_n/\varrho'(x_n)])]_{\mathbf{T}}$$

and

$$[(\lambda x.M[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})])]_{\mathbf{T}} = [(\lambda x.N[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})])]_{\mathbf{T}}$$

and so

$$\|(\lambda x.M)\|^{\varrho'} = \|(\lambda x.N)\|^{\varrho'}\,.$$

But x is not free in $(\lambda x.M)$ or $(\lambda x.N)$, so $\|(\lambda x.M)\|^{\varrho} = \|(\lambda x.N)\|^{\varrho}$. $\square$

Note that we cannot carry this proof through in the case of the closed term model because $[x]_{\mathbf{T}^0}$ is not defined. By Theorem 1 we have the following

Corollary. *The open term model is a syntactical λ–algebra.* □

Next we want

Theorem 3. *The closed term algebra is a syntactical λ-algebra.*

Proof. It suffices to show

$$\mathbf{T} \vdash M = N \implies \mathcal{M}^0(\mathbf{T}) \models M = N.$$

So suppose $\mathbf{T} \vdash M = N$, and $\mathbf{FV}(M) \cup \mathbf{FV}(N) \subseteq \{x_1, \ldots, x_n\}$. Let $\varrho : \mathbf{Var} \to X^0$ be an arbitrary valuation. Then

$$\mathbf{T} \vdash M[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})] = N[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})]$$

$$[M[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n)]]_{\mathbf{T}^0} = [N[x_1/\varrho(x_1), \ldots, x_n/\varrho(x_n)]]_{\mathbf{T}^0}$$

$$\|M\|^{\varrho} = \|N\|^{\varrho}.$$

Since ϱ is arbitrarily selected, $\mathcal{M}^0(\mathbf{T}) \models M = N$. □

We commented above that the proof that the open term model is a λ-model cannot be used to prove the same thing for the closed term model. The next theorem show that this is not an accidental feature of the proof we gave.

Theorem 4. *The closed term model is not a λ–model.*

To prove this we need the following lemma.

Lemma 4 (Existence of Plotkin terms). *There exist terms $\Phi, \Psi \in \Lambda^0(\emptyset)$ such that*

(1) *For all $Z \in \Lambda^0(\emptyset)$, $\vdash \Phi(Z) = \Psi(Z)$;*

(2) *for $x \in \mathbf{Var}$, $\nvdash \Phi(x) = \Psi(x)$.*

Proof sketch for lemma. We satisfy ourselves with a sketch, since many tedious details are required for a proper proof—the sketch will no doubt strike most readers as plenty long enough. We begin by noting that we can use Turing's fixpoint combinator to prove the *Double Fixpoint Theorem*: For any pair of combinators F, G, there is a pair of combinators A, B such that $A = F(A)(B)$ and $B = G(A)(B)$ (see [Barendregt 1984, 6.5.1]).

We need some terminology. So we will make the following definitions.

(1) *Finite sequences*: $\langle M_0, \ldots, M_n \rangle \equiv (\ldots(\lambda z.z(M_0)(M_1)\ldots(M_n))$. (Though we won't need them, it is probably worth mentioning that the projections can be defined by $\pi_i^n \equiv (\lambda x.x[\lambda x_0(\ldots(\lambda x_n.x_i)\ldots)]))$.

(2) *Numerals*. First put $\mathbf{T} \equiv (\lambda x.(\lambda y.x))$ and $\mathbf{F} = (\lambda x.(\lambda y.y))$. Now we define

$$\ulcorner 0 \urcorner \equiv (\lambda x.x) \qquad \ulcorner n+1 \urcorner \equiv \langle \mathbf{F}, \ulcorner n \urcorner \rangle$$

$$\mathbf{S} \equiv (\lambda x. \langle \mathbf{F}, x \rangle) \qquad \mathbf{P} \equiv (\lambda x.x(\mathbf{F})) \qquad \mathbf{Zero} \equiv (\lambda x.x(\mathbf{T})).$$

It is then easily shown that

$$\mathbf{Zero}(\ulcorner 0 \urcorner) = \mathbf{T} \qquad \mathbf{Zero}(\ulcorner n+1 \urcorner) = \mathbf{F}$$

$$\mathbf{S}(\ulcorner n \urcorner) = \ulcorner n+1 \urcorner \qquad \mathbf{P}(\ulcorner n+1 \urcorner) = \ulcorner n \urcorner$$

and that the numerals are all distinct normal forms. We thus have defined what is often called a *normal numeral system*.

(3) *Gödel numbering*: It is straightforward to define in one of the standard ways a map $\# : \Lambda(\emptyset) \to \mathbb{N}$ where $\mathbb{N}$ is the set of natural numbers. We write for each $M \in \Lambda(\emptyset)$, hazarding the ambiguity, '$\ulcorner M \urcorner$' for '$\ulcorner \# M \urcorner$'.

It is then possible to define a closed term $\mathbf{E} \in \Lambda^0(\emptyset)$ such that, for every $M \in \Lambda^0(\emptyset)$, $\vdash \mathbf{E}(\ulcorner M \urcorner) \triangleright M$ (see [Barendregt 1984, 8.1.6]). The reader will recall from §3.C that it follows from this that $\vdash \mathbf{E}(\ulcorner M \urcorner) = M$.

Now we put

$$A \equiv (\lambda f.(\lambda g.(\lambda x.(\lambda y.(\lambda z.f(x)[f(\mathbf{S}(x))(g(\mathbf{S}(x)))(z)(y)](\mathbf{E}(x))))))).$$

$$B \equiv (\lambda f.(\lambda g.(\lambda x.f(\mathbf{S}(x))(g(\mathbf{S}(x)))(\mathbf{E}(\mathbf{S}(x)))(g(x))))).$$

The Double Fixpoint Theorem gives us combinators F, and G such that $A(F)(G) = F$ and $B(F)(G) = G$. These fixpoints give us

$$(*) \qquad F(x)(y)(z) = F(x)[F(\mathbf{S}(x))(G(\mathbf{S}(x)))(z)(y)](\mathbf{E}(x))$$

$$(**) \qquad G(x) = F(\mathbf{S}(x))(G(\mathbf{S}(x)))(\mathbf{E}(\mathbf{S}(x)))(G(x))$$

Definition 4 (Plotkin terms). Let F and G be the combinators just described. Now we define $\Phi \equiv F(\ulcorner 0 \urcorner)(G(\ulcorner 0 \urcorner))$ and $\Psi \equiv (\lambda x.\Phi(\lambda y.y))$.

The reason we have gone to such trouble is to be able to show that any $m, n \in \mathbb{N}$,

$$(\dagger) \qquad \vdash F(\ulcorner n \urcorner)(G(\ulcorner n \urcorner))(\mathbf{E}(\ulcorner n \urcorner)) = F(\ulcorner n \urcorner)(G(\ulcorner n \urcorner))(\mathbf{E}(\ulcorner n+m \urcorner)).$$

We show this by induction on m. If $m = 0$, this is trivial. We have, for $m > 0$,

$$\begin{aligned}
&F(\ulcorner n \urcorner)(G(\ulcorner n \urcorner))(\mathbf{E}(\ulcorner n \urcorner)) \\
&\quad = F(\ulcorner n \urcorner)[F(\ulcorner n+1 \urcorner)(G(\ulcorner n+1 \urcorner))(\mathbf{E}(\ulcorner n+1 \urcorner))(G(n))](\mathbf{E}(\ulcorner n \urcorner)) \\
&\quad = F(\ulcorner n \urcorner)[F(\ulcorner n+1 \urcorner)(G(\ulcorner n+1 \urcorner)(\mathbf{E}(\ulcorner (n+1)+(m-1) \urcorner))(G(\ulcorner n \urcorner)](\mathbf{E}(\ulcorner n \urcorner)) \\
&\quad = F(\ulcorner n \urcorner)(G(\ulcorner n \urcorner))(\mathbf{E}(\ulcorner n+m \urcorner))
\end{aligned}$$

by $**$, IH, and $*$ respectively.

We can now show that (1) holds. By (†) we have $\Phi(\mathbf{E}(\ulcorner 0 \urcorner)) = \Phi(\mathbf{E}(\ulcorner m \urcorner))$ for any $m \in \mathbb{N}$. So we have that for every $Z, P, \in \Lambda^0(\emptyset)$

$$\Phi(\mathbf{E}(\ulcorner Z \urcorner)) = \Phi(\mathbf{E}(\ulcorner P \urcorner)) = \Phi(\mathbf{E}(\ulcorner 0 \urcorner))$$

and so by the feature of $\mathbf{E}$ mentioned above we have

$$\Phi(Z) = \Phi(P) = \Phi(\mathbf{E}(\ulcorner 0 \urcorner)).$$

So for every $Z \in \Lambda^0(\emptyset)$

$$\Phi(Z) = \Phi((\lambda y.y)) = \Psi(Z).$$

As for (2), it is messy but not hard to show that for any $x \in \mathbf{Var}, \vdash \Phi(x) \triangleright M \implies x \in \mathbf{FV}(M)$; it is also clear by inspection of the rules and axioms for $\triangleright$ that if $\vdash M \triangleright N$ then $\mathbf{FV}(N) \subseteq \mathbf{FV}(M)$. We can use these facts to show that $\nvdash \Phi = \Psi$. For suppose $\vdash \Phi = \Psi$. Then $\Phi(x) = \Psi(x) = \Phi(\lambda y.y)$. By the Church-Rosser Theorem, this gives us both $\vdash \Phi(x) \triangleright M$ and $\vdash \Phi((\lambda y.y)) \triangleright M$, and so we have both $x \in \mathbf{FV}(M)$ and $x \notin \mathbf{FV}(M)$, a contradiction. $\square$

At last we can use Lemma 4 to give the

Proof of Theorem 4. Let $\varrho : \mathbf{Var} \to X^0$. First, by our definitions of interpretations in the closed term model and by Lemma 4 we have

$$\begin{aligned} \|\Phi(x)\|^{\varrho} = \|\Phi(x)[x/\varrho(\mathbf{x})]\| = [\Phi(x)[x/\varrho(x)]]_{\emptyset} \\ = [\Psi(x)[x/\varrho(x)]]_{\emptyset} = \|\Psi(x)[x/\varrho(\mathbf{x})]\| = \|\Psi(x)\|^{\varrho}. \end{aligned}$$

However,

$$\mathcal{M}^0(\emptyset) \models (\lambda x.\Phi(x)) = (\lambda x.\Psi(x)) \iff \|(\lambda x.\Phi(x))\|^{\varrho} = \|(\lambda x.\Psi(x))\|^{\varrho} \qquad \text{for all } \varrho$$

$$\iff \vdash (\lambda x.\Phi(x)) = (\lambda x.\Psi(x))$$

$$\iff \vdash \Phi(x) = \Psi(x).$$

Since x is the only free variable in $\Phi(x)$ and $\Psi(x)$, we have $\|\Phi(x)\|^{\varrho(x/a)} = \|\Psi(x)\|^{\varrho(x/a)}$, for all a, but $\|(\lambda x.\Phi(x))\|^{\varrho} \neq \|(\lambda x.\Psi(x))\|^{\varrho}$. □

This theorem tells us that the λ-calculus is ω-incomplete in a sense we can make precise by stating the

Corollary. *The rule*

$$\frac{M(P) = N(P)}{M = N} \qquad \text{for all closed } P$$

is not a valid rule of inference in the λ–calculus. □

At last we can state

Theorem 5 (Completeness Theorems).

(1) $\vdash M = N \iff M = N$ *is satisfied in all syntactical λ-models (or syntactical λ-algebras).*

(2) *Let* $\mathbf{T}$ *be a set of closed λ-equations.*

$\mathbf{T} \vdash M = N \iff M = N$ *is satisfied in all syntactical λ-models satisfying* $\mathbf{T}$.

(3) *Let* $\mathbf{T}$ *be a set of closed λ-equations.*

$\mathbf{T} \vdash M = N \iff M = N$ *is satisfied in all syntactical λ-algebras satisfying* $\mathbf{T}$, *provided M and N are closed.*

Proof.

(1) ($\Longrightarrow$) is the soundness lemma. For ($\Longleftarrow$), if $M = N$ is satisfied in all models (algebras), $\mathcal{M}(\emptyset) \models M = N \iff [M]_\emptyset = [N]_\emptyset \iff \vdash M = N$.

(2) Same with $\mathcal{M}(\mathbf{T})$ for $\mathcal{M}(\emptyset)$.

(3) Ditto. □

We note however that we cannot prove an analogue for (2) for syntactical λ-algebras. This follows from the fact that it is possible for $M = N$ to be satisfied in all syntactical λ-algebras satisfying $\mathbf{T}$ and yet $\mathbf{T} \nvdash M = N$. This is another instance of ω-incompleteness. For a proof, see [Barendregt 1984, pp. 97-98]. See §5 for some discussion of this.

2. First Order Semantics

The definition of interpretations given in §1 is rather transparently just a restatement of the syntactical rules in semantical form. While this has allowed us to prove useful facts about the λ-calculus, it has left the nature of the structures themselves rather a mystery. We will now give a more mathematical characterization of the λ-models and λ-algebras so that we can begin to eliminate this mystery.

Recall that $\mathcal{M} = \langle X, \circ \rangle$ is an *applicative structure* if $\circ$ is a binary operation on X. We again use boldface to indicate the name for the elements of X. We will define the *set of terms over* $\mathcal{M}$, which we denote by '$T(\mathcal{M})$', to be the smallest set such that:

$$\mathbf{Var} \subseteq T(\mathcal{M})$$

$$B \in \mathcal{M} \implies \mathbf{B} \in T(\mathcal{M})$$

$$A, B \in \mathcal{M} \implies A(B) \in T(\mathcal{M}).$$

We will use 'A', 'B', ... to denote arbitrary terms of $T(\mathcal{M})$, and we write as usual '$B \in \mathcal{M}$' for '$B \in X$'.

Obviously we're not interested in all applicative structures. There are two equivalent ways of specifying the sort of applicative structure which will be of interest to us. We now define both.

Definition 5. A *combinatory algebra* is an applicative structure with distinguished elements k and s, $\mathcal{M} = \langle X, \circ, k, s \rangle$, such that for any $a, b, c \in X$

(1) $(k \circ a) \circ b = a$

(2) $(((s \circ a) \circ b) \circ c) = ((a \circ c) \circ (b \circ c))$.

A *valuation* is a map $\varrho : \mathbf{Var} \to X$. We define the *interpretation of a term* $A \in T(\mathcal{M})$ *in* $\mathcal{M}$ *under* ϱ, denoted by '$\|A\|^{\varrho}_{\mathcal{M}}$', inductively as follows:

$$\|x\|^{\varrho}_{\mathcal{M}} = \varrho(x)$$

$$\|\mathbf{B}\|^{\varrho}_{\mathcal{M}} = B$$

$$\|A(B)\|^{\varrho}_{\mathcal{M}} = \|A\|^{\varrho}_{\mathcal{M}} \circ \|B\|^{\varrho}_{\mathcal{M}} .$$

Definition 6. An applicative structure $\mathcal{M}$ is *combinatory complete* if and only if for every $A \in T(\mathcal{M})$ with $\mathbf{FV}(A) \subseteq \{x_1, \ldots, x_n\}$ there is an $a \in X$ such that for any $b_1, \ldots, b_n \in X$

$$(\ldots((a \circ b_1) \circ b_2) \circ \cdots \circ b_n) = \|A\|_{\mathcal{M}}^{\varrho(x_1/b_1,\ldots,x_n/b_n)} .$$

It is not difficult to prove the following result, which goes back to the 1924 work of Schönfinkel. A proof can be found in any of the usual references.

Theorem 6. *An applicative structure is combinatory complete if and only if it can be expanded to a combinatory algebra by choosing k and s.* □

It is useful to have both these definitions in front of us, since the definition in terms of combinatory completeness obviously is independent of our syntax, while combinatory algebras are easier to work with. For interest's sake, we digress to quote the following proposition.

Proposition 1. *Non-trivial combinatory algebras are never commutative, never associative, never finite, and never recursive.* □

For a proof see [Koymans p. 58] or [Barendregt 1984 pp. 91-2]. This proposition shows just how unalgebraic these algebras are.

The reader is no doubt already in a position to guess the definition of interpretation for most terms of the λ-calculus in combinatory algebras, but we so far have no way to interpret the λ-abstraction terms. To rectify this we first extend $T(\mathcal{M})$ by adding new constants $\mathbf{K}$ and $\mathbf{S}$, which denote k and s respectively. We also define $\mathbf{I} = (\mathbf{S} \circ \mathbf{K}) \circ \mathbf{K}$. Now for $A \in T(\mathcal{M})$ and $x \in \mathbf{Var}$, we define $(\lambda^* x.A) \in T(\mathcal{M})$

inductively by

$$(\lambda^* x.x) = \mathbf{I}$$

$$(\lambda^* x.P) = \mathbf{K} \circ P \qquad \text{if } P \text{ does not contain } x$$

$$(\lambda^* x.P(Q)) = (\mathbf{S} \circ (\lambda^* x.P)) \circ (\lambda^* x.Q).$$

We will need the following notation: if $\mathcal{M}$ is an applicative structure, then '$\Lambda(\mathcal{M})$' is used to denote $\Lambda(\{\, \mathbf{A} \mid A \in \mathcal{M} \,\})$, i. e. the set of λ-terms we get if we take the names of members of X as constants. Note that together with the set of terms over $\mathcal{M}$ defined at the start of this section this gives us two sets of terms to consider. (Readers puzzled by the restriction on P in the second clause of the above definition are reminded of the simple definition of the terms over $\mathcal{M}$, which ensures that if x is in P it must fall under one of the other clauses.)

Next we define two maps

$$\mathbf{CL} : \Lambda(\mathcal{M}) \to T(\mathcal{M})$$

$$\mathbf{lam} : T(\mathcal{M}) \to \Lambda(\mathcal{M})$$

as follows:

$$\mathbf{CL}(x) = x$$

$$\mathbf{CL}(\mathbf{A}) = \mathbf{A}$$

$$\mathbf{CL}(M(N)) = \mathbf{CL}(M) \circ \mathbf{CL}(N)$$

$$\mathbf{CL}(\lambda x.M) = (\lambda^* x.\mathbf{CL}(M))$$

$$\mathbf{lam}(x) = x$$

$$\mathbf{lam}(\mathbf{A}) = \mathbf{A}$$

$$\mathbf{lam}(A \circ B) = \mathbf{lam}(A)(\mathbf{lam}(B))$$

$$\mathbf{lam}(\mathbf{K}) = (\lambda x.(\lambda y.x))$$

$$\mathbf{lam}(\mathbf{S}) = (\lambda x.(\lambda y.(\lambda z.x(z)(y(z))))).$$

Finally we are in a position to define a map which interprets the λ-terms in combinatory algebras. For $M \in \Lambda(\mathcal{M})$, the interpretation of M in $\mathcal{M}$ under the valuation ϱ, denoted by '$\|M\|^{\varrho}_{\mathcal{M}}$', is defined by

$$\|M\|^{\varrho}_{\mathcal{M}} = \|\mathbf{CL}(M)\|^{\varrho}_{\mathcal{M}} .$$

As usual we will put

$$\mathcal{M}, \varrho \models M = N \iff \|M\|^{\varrho}_{\mathcal{M}} = \|N\|^{\varrho}_{\mathcal{M}}$$

$$\mathcal{M} \models M = N \iff \mathcal{M}, \varrho \models M = N \qquad \text{for all } \varrho.$$

However, we are not quite home yet. There are equations provable in λ-calculus which are not true in every combinatory algebra. For example, there are combinatory algebras such that

$$\mathcal{M} \not\models (\lambda z.(\lambda x.x)(z)) = (\lambda z.z).$$

A proof of this can be found in [Barendregt 1984, p. 93]. So we must make the following

Definition 7. A combinatory algebra $\mathcal{M}$ is a λ-*algebra* if and only if for all $A, B \in T(\mathcal{M})$,

$$\vdash \mathbf{lam}(A) = \mathbf{lam}(B) \implies \mathcal{M} \models A = B.$$

We now state a few useful facts for use below.

Comments. First we note that by an easy inductive argument we can show that $(\lambda^* x.M)(x) = M$ (remember that the variable conventions assure us that $x \notin$

$\mathbf{FV}(M)$). We can use this fact to show that

$$\mathbf{CL}(\mathbf{lam}(A)) = A.$$

This is shown by an inductive argument on the complexity of A. The variable and constant cases are trivial. If $A \equiv \mathbf{S}$, $\mathbf{CL}(\mathbf{lam}(\mathbf{S})) = (\lambda^* x.(\lambda^* y.(\lambda^* z.x(z)(y(z))))) = (\lambda^* x.(\lambda^* y.\mathbf{S}((\lambda^* z.x(z)))((\lambda^* z.y(z)))))$, which is $(\lambda^* x.(\lambda^* y.\mathbf{S}(x(y)))) = \mathbf{S}$ by repeated application of the above fact. Similarly, $\mathbf{CL}(\mathbf{lam}(\mathbf{K})) = (\lambda^* x.(\lambda^* y.x)) = (\lambda^* x.\mathbf{K}(x))$ by the definition of λ^*, and this is $\mathbf{K}$ by the fact we started with.

Finally, we can show that for $M \in \Lambda(\mathcal{M})$, $\vdash \mathbf{lam}(\mathbf{CL}(M)) = M$. The proof is by induction on the complexity of M, and it is entirely routine if it is a variable, constant, or term of form $P(Q)$. For the case of abstraction terms consider $M \equiv (\lambda x.P)$ and suppose the claim is true for all terms with complexity less than that of M. By the definition of $\mathbf{CL}$, $\mathbf{lam}(\mathbf{CL}(M)) = \mathbf{lam}(\lambda^* x.\mathbf{CL}(P))$, and by the definition of λ^* there are three cases to consider. First,

$$\mathbf{lam}(\mathbf{I}) \equiv \mathbf{lam}((\mathbf{S} \circ \mathbf{K}) \circ \mathbf{K}) \equiv [\mathbf{lam}(\mathbf{S})(\mathbf{lam}(\mathbf{K}))](\mathbf{lam}(\mathbf{K}))$$

which some calculation shows to be $(\lambda z.z)$. The result follows by α-convertibility. If $\mathbf{CL}(P) \equiv Q$ with x not in Q we have

$$\mathbf{lam}(\mathbf{CL}(P)) = \mathbf{lam}(\mathbf{K} \circ Q) = \mathbf{lam}(\mathbf{K})[\mathbf{lam}(Q)] = (\lambda x.(\lambda y.x))(P) = (\lambda y.P).$$

Finally, suppose $\lambda^* x.P = (\mathbf{S} \circ \lambda^* x.R) \circ \lambda^* x.Q$. This is $(\mathbf{S} \circ \mathbf{CL}(\lambda x.R)) \circ \mathbf{CL}(\lambda x.Q)$, so $\mathbf{lam}(\mathbf{CL}(M)) = (\lambda z.\mathbf{lam}(\mathbf{CL}((\lambda x.P)))(z)(\mathbf{lam}(\mathbf{CL}((\lambda x.Q)))(z))$ which some calculation will show to be $(\lambda z.\mathbf{lam}(\lambda^* x.\mathbf{CL}(P))(z)((\lambda^* x.\mathbf{CL}(Q))(z))$ and by the above facts $\lambda z.\mathbf{lam}(\mathbf{CL}(P(Q))$ and by inductive hypothesis we have $(\lambda z.P(Q))$.

With these facts in our pockets it is easy to prove

Lemma 5. *Let $\mathcal{M}$ be a combinatory algebra. Then $\mathcal{M}$ is a λ-algebra if and only if for all $M, N \in \Lambda(\mathcal{M})$*

(1) $\vdash M = N \implies \mathcal{M} \models M = N$

(2) $\mathcal{M} \models \mathbf{CL}(\mathbf{lam}(\mathbf{K})) = \mathbf{K}$ *and* $\mathcal{M} \models \mathbf{CL}(\mathbf{lam}(\mathbf{S})) = \mathbf{S}$.

Proof. ($\Longrightarrow$) By the last remark

$$\begin{aligned} \vdash M = N &\implies \vdash \mathbf{lam}(\mathbf{CL}(M)) = \mathbf{lam}(\mathbf{CL}(N)) \\ &\implies \mathcal{M} \models \mathbf{CL}(M) = \mathbf{CL}(N) && \text{since } \mathcal{M} \text{ is a } \lambda\text{-algebra} \\ &\implies \mathcal{M} \models M = N && \text{by def'n of } \models. \end{aligned}$$

Furthermore, for all $A \in T(\mathcal{M})$ we have $\vdash \mathbf{lam}(\mathbf{CL}(\mathbf{lam}(A))) = \mathbf{lam}(A)$ so $\mathcal{M} \models \mathbf{CL}(\mathbf{lam}(A)) = A$.

($\Longleftarrow$) By induction on $A \in T(\mathcal{M})$, (2) gives us $\mathcal{M} \models \mathbf{CL}(\mathbf{lam}(A)) = A$. So by (1),

$$\begin{aligned} \vdash \mathbf{lam}(A) = \mathbf{lam}(B) &\implies \mathcal{M} \models \mathbf{CL}(\mathbf{lam}(A)) = \mathbf{CL}(\mathbf{lam}(B)) \\ &\implies \mathcal{M} \models A = B. \quad \Box \end{aligned}$$

We now introduce the familiar notion of weak extensionality to our new semantics. Let $\mathcal{M}$ be a combinatory algebra. $\mathcal{M}$ is said to be *weakly extensional* if and only if

$$\|M\|_{\mathcal{M}}^{\varrho(x/a)} = \|N\|_{\mathcal{M}}^{\varrho(x/a)} \quad \text{for all } a \in \mathcal{M} \implies \mathcal{M} \models (\lambda x.M) = (\lambda x.N).$$

Obviously we want to introduce a notion of λ-model which is equivalent to "weakly extensional λ-algebra", but we want to do so in a way that is less tied to our syntax, and so is hopefully more informative. The usual way of doing this is due to [Scott 1980b]. First we put $\mathbf{1} \equiv \mathbf{S}(\mathbf{K}(\mathbf{I}))$, then we make

Definition 8. A *λ-model* is a λ-algebra $\mathcal{M}$ such that for any $a, b \in \mathcal{M}$,

$$a \circ x = b \circ x \text{ for every } x \in \mathcal{M} \implies \mathbf{1} \circ a = \mathbf{1} \circ b.$$

Remarks. We will take this opportunity to collect together some useful facts about the special combinators we have defined. First, the reader should have no trouble becoming convinced that **I** is aptly named, since $\mathbf{I}(a) = a$ for any a in any combinatory algebra. Now $(\mathbf{1}\circ a)\circ b = (\mathbf{S}\circ(\mathbf{K}(\mathbf{I}))\circ a)\circ b = \mathbf{K}(\mathbf{I})(b)\circ(a\circ b) = \mathbf{I}\circ(a\circ b)$. So

(i) $$(\mathbf{1} \circ a) \circ b = a \circ b.$$

If $\mathcal{M}$ is a λ-algebra, we can reason as follows for our distinguished elements **S** and **K**. In proving the last lemma we showed that $\mathbf{lam}(\mathbf{CL}(\mathbf{lam}(\mathbf{S}))) = \mathbf{lam}(\mathbf{S})$. Since $\mathcal{M}$ is a λ–algebra, we know $\mathcal{M} \models \mathbf{CL}(\mathbf{lam}(\mathbf{S})) = \mathbf{S}$. By definition, $\|\mathbf{S}\| = \mathbf{S}$ and $\|\mathbf{CL}(\mathbf{lam}(\mathbf{S}))\|_{\mathcal{M}} = \|\mathbf{lam}(\mathbf{S})\|_{\mathcal{M}}$, so $\mathbf{S} = \|\mathbf{lam}(\mathbf{S})\|$. Similarly for **K**, and so also for **I** and **1**. So $\mathbf{1} = \mathbf{S}(\mathbf{K}(\mathbf{I})) = \|\mathbf{lam}(\mathbf{S})(\mathbf{lam}(\mathbf{K})(\mathbf{lam}(\mathbf{I})))\|_{\mathcal{M}}$, which a little calculation will show to be $\|(\lambda y.(\lambda z.y(z))\|_{\mathcal{M}}$. So, recalling that we have a constant **a** in $\Lambda(\mathcal{M})$ for each $a \in \mathcal{M}$, we have

(ii) $$\mathbf{1} \circ a = \|(\lambda y.\mathbf{a}(y))\|_{\mathcal{M}},$$

from which the following facts easily follow:

(iii) $$\mathbf{1} \circ \|(\lambda x.A)\|_{\mathcal{M}} = \|(\lambda x.A)\|_{\mathcal{M}} \qquad \text{for all } A \in T(\mathcal{M})$$

(iv) $$\mathbf{1} \circ \mathbf{1} = \mathbf{1}.$$

These facts will be useful for proving

Theorem 7. *$\mathcal{M}$ is a λ-model $\iff$ $\mathcal{M}$ is a weakly extensional λ-algebra.*

Proof. Suppose $\mathcal{M}$ is weakly extensional. Then if, for every $c \in \mathcal{M}$, $a \circ c = b \circ c$ we have $\|\mathbf{a}(x)\|_{\mathcal{M}}^{\varrho(x/c)} = \|\mathbf{b}(x)\|_{\mathcal{M}}^{\varrho(x/c)}$ for all c, and so $\mathcal{M} \models (\lambda x.a(x)) = (\lambda x.b(x))$. So $\mathbf{1} \circ a = \mathbf{1} \circ b$, by (ii).

So suppose $\mathcal{M}$ is a λ-model and that $\|A\|_{\mathcal{M}}^{\varrho(x/c)} = \|B\|_{\mathcal{M}}^{\varrho(x/c)}$ for all $c \in \mathcal{M}$. Since $\|A\|_{\mathcal{M}}^{\varrho(x/c)} = \|A[x/\mathbf{c}]\|_{\mathcal{M}}^{\varrho} = \|(\lambda x.A)(\mathbf{c})\|_{\mathcal{M}}^{\varrho}$ and similarly for B, we have

$$\|(\lambda x.A)\|_{\mathcal{M}}^{\varrho} \circ c = \|(\lambda x.B)\|_{\mathcal{M}}^{\varrho} \circ c \text{ for all } c \in \mathcal{M}.$$

By the definition of λ-model, this gives us

$$\mathbf{1} \circ \|(\lambda x.A)\|_{\mathcal{M}}^{\varrho} = \mathbf{1} \circ \|(\lambda x.B)\|_{\mathcal{M}}^{\varrho}$$

and so $\|(\lambda x.A)\|_{\mathcal{M}}^{\varrho} = \|(\lambda x.B)\|_{\mathcal{M}}^{\varrho}$ by (iii). $\square$

We will digress to state the following proposition, though we won't bother to prove it. An applicative structure $\mathcal{M}$ is said to be *extensional* if and only if for every $a, b \in \mathcal{M}$

$$a \circ c = b \circ c \quad \text{for all} \quad c \in \mathcal{M} \implies a = b.$$

Proposition 2. *If $\mathcal{M}$ is a λ-algebra, then $\mathcal{M}$ is extensional if and only if $\mathcal{M}$ is weakly extensional and $\mathcal{M} \models \mathbf{I} = \mathbf{1}$.* $\square$

To show that this description of the models for λ-calculus is equivalent to the syntactical description given earlier, it will be useful to consider both types of structures as categories. We will begin with the λ-algebras.

Definition 9. Let $\mathcal{M}_1 = \langle X_1, \circ_1, k_1, s_1 \rangle$ and $\mathcal{M}_2 = \langle X_2, \circ_2, k_2, s_2 \rangle$ be two combinatory algebras. Then $\varphi : X_1 \to X_2$ is a *(combinatory algebra) homomorphism* if and only if the three equations

$$\varphi(x \circ_1 y) = \varphi(x) \circ_2 \varphi(y)$$

$$\varphi(k_1) = k_2$$

$$\varphi(s_1) = s_2$$

hold. We say $\mathcal{M}_1$ *is homomorphic to* $\mathcal{M}_2$ if there is a homomorphism $\varphi : X_1 \to X_2$, and sometimes write '$\varphi : \mathcal{M}_1 \to \mathcal{M}_2$' for this. $\mathcal{M}_1$ is *embeddable* in $\mathcal{M}_2$ if $\varphi : \mathcal{M}_1 \to \mathcal{M}_2$ for an injective φ. We write '$\mathcal{M}_1 \preccurlyeq \mathcal{M}_2$' for this. If φ is the identity, we say $\mathcal{M}_1$ is a *substructure* of $\mathcal{M}_2$. We say $\mathcal{M}_1$ is *isomorphic* to $\mathcal{M}_2$ (and we write '$\mathcal{M}_1 \cong \mathcal{M}_2$') if '$\varphi : \mathcal{M}_1 \to \mathcal{M}_2$' for a bijective φ.

It is not hard to show that we get a category if we take as objects the λ algebras and as arrows the combinatory algebra homomorphisms between them. We get our other category by taking as objects the syntactical λ-algebras, and as arrows the homomorphisms between them, which are defined like this.

Definition 10. For any two syntactical λ-algebras $\mathcal{M}_1 = \langle X_1, \circ_1, \|\cdot\| \rangle$ and $\mathcal{M}_2 = \langle X_2, \circ_2, \|\cdot\| \rangle$, a *(syntactical λ-algebra) homomorphism* is a map $\varphi : X_1 \to X_2$

such that

$$\varphi\left(\|M\|^{\varrho}_{\mathcal{M}_1}\right) = \|\varphi(M)\|^{\varphi\cdot\varrho}_{\mathcal{M}_2}$$

where we replace any of the constants $\mathbf{a}$ in M by the constant $\varphi(\mathbf{a})$ in $\varphi(M)$.

This is also easily seen to be a category. We can now show the following.

Theorem 8. *The category of syntactical λ-algebras and homomorphisms and the category of λ-algebras and homomorphisms are isomorphic. Moreover, syntactical λ-models correspond exactly to λ-models under this isomorphism.*

Proof. For a syntactical λ-algebra $\mathcal{M} = \langle X, \circ, \|\cdot\|\rangle$, define

$$F(\mathcal{M}) = \langle X, \circ, \|\mathbf{K}\|_{\mathcal{M}}, \|\mathbf{S}\|_{\mathcal{M}}\rangle\,.$$

This is easily shown to be a combinatory algebra using definition 1. For instance, $(\|(\lambda x.(\lambda y.x))\|_{\mathcal{M}} \circ a) \circ b = \|x\|^{\varrho(x/a)(y/b)}_{\mathcal{M}} = a$ for any ϱ, so the axiom for k is satisfied. An easy inductive argument (with the usual calculations required for the abstraction term step) shows that the following trivial looking fact holds: $\|M\|^{\varrho}_{\mathcal{M}} = \|M\|^{\varrho}_{\mathcal{M}}$. What makes this non-trivial is that the expression on the right refers to the interpretation in the algebras, while the expression on the left refers to interpretation in the syntactical structures. So the Soundness Theorem tells us that $\vdash M = N \implies \mathcal{M} \models M = N$. By our current definition of $\mathbf{K}$ we have $\mathbf{lam}(\mathbf{K}) = \mathbf{K}$, so $\|\mathbf{CL}(\mathbf{lam}(\mathbf{K}))\|^{\varrho}_{\mathcal{M}} = \|\mathbf{CL}(\mathbf{K})\|^{\varrho}_{\mathcal{M}} = \|\mathbf{K}\|^{\varrho}_{\mathcal{M}}$ for any ϱ, and similarly for $\mathbf{S}$. So $\mathcal{M}$ is a λ-algebra.

For $\varphi : \mathcal{M}_1 \to \mathcal{M}_2$, define

$$F\varphi = \varphi : F(\mathcal{M}_1) \to F(\mathcal{M}_2).$$

It is routine to show that this is a combinatory algebra homomorphism.

Conversely, for a λ-algebra $\mathcal{A} = \langle X, \circ, k, s\rangle$ define

$$G(\mathcal{A}) = \langle X, \circ, \|\cdot\|_{\mathcal{A}}\rangle$$

and for $\varphi : \mathcal{A}_1 \to \mathcal{A}_2$ define

$$G(\varphi) = \varphi : G(\mathcal{A}_1) \to G(\mathcal{A}_2).$$

To show that $G(\mathcal{A})$ is a syntactical λ–algebra requires that we show that $\|\cdot\|_{\mathcal{A}}$ gives us an interpretation, which is straightforward, and that $\vdash M = N \implies \mathcal{M} \models M = N$. The second requires again that we show $\|M\|^{\varrho}_{\mathcal{M}} = \|M\|^{\varrho}_{\mathcal{M}}$ where each side of the equation denotes interpretation is a different sense.

To complete the proof we only need to observe that F is obviously a functor and G is an inverse for it. $\square$

This theorem is quite useful. It allows us to simply speak of a λ-model or λ-algebra, with no need to specify whether we intend a *syntactical* model or algebra or not. Furthermore, it allows us to simply carry over the results from our discussion of syntactical models and algebras to the new context. We only need to translate via the map F. In particular, for any λ-theory $\mathbf{T}$ the term models for $\mathbf{T}$ will be $\mathcal{M}(\mathbf{T}) = \langle X, \bullet, [\mathbf{K}]_{\mathbf{T}}, [\mathbf{S}]_{\mathbf{T}}\rangle$ and $\mathcal{M}^0(\mathbf{T}) = \left\langle X^0, \bullet^0, [\mathbf{K}]_{\mathbf{T}^0}, [\mathbf{S}]_{\mathbf{T}^0}\right\rangle$ (cf. definition 3). So we have, with no further ado,

Corollary. *For any λ–theory $\mathbf{T}$,*

(1) $\mathcal{M}^0(\mathbf{T})$ *is a λ-algebra.*

(2) *$\mathcal{M}(\mathbf{T})$ is a λ-model.*

(3) *$\mathcal{M}^0(\mathbf{T})$ is a not, in general, a λ-model.* □

We are now in a position to say some interesting things about the relationship between the λ-models and the λ-algebras. We begin with some terminology. Let $\mathcal{A}$ be a combinatory algebra. The *interior* of $\mathcal{A}$ (denoted by '$\mathcal{A}^0$') is the substructure of $\mathcal{A}$ generated by k and s. We say that $\mathcal{A}$ is *hard* if $\mathcal{A}^0 = \mathcal{A}$. We will use '$T'(\mathcal{A})$' to denote the set of $M \in T(\mathcal{A})$ such that M contains only $\mathbf{K}$, $\mathbf{S}$ and variables. We now define two sets:

$$Th(\mathcal{A}) = \{\, M = N \mid \mathcal{A} \models M = N, M, N \in T'(\mathcal{A}), M,\ N \text{ are closed} \,\}$$

$$Th(\overline{\mathcal{A}}) = \{\, M = N \mid \mathcal{A} \models M = N, M, N \in T(\mathcal{A}), M,\ N \text{ are closed} \,\}$$

Obviously $Th(\mathcal{A}) \subseteq Th(\overline{\mathcal{A}})$. It is not hard to see that, up to isomorphism, the closed term model is the interior of the open term model for any theory $\mathbf{T}$.

Proposition 3. *Let $\mathcal{A}$ be a λ-algebra. Then*

(1) $\mathcal{M}^0(Th(\mathcal{A})) \cong \mathcal{A}^0$.

(2) $\mathcal{M}^0(Th(\overline{\mathcal{A}})) \cong \mathcal{A}$.

Proof.

(1) For $M, N \in T'(\mathcal{A})$, we have $\mathcal{A} \models M = N \iff M = N \in Th(\mathcal{A}) \iff N \in [M]_{Th(\mathcal{A})}$. So $\varphi([M]_{Th(\mathcal{A})}) = \|M\|$ defines an isomorphism as required.

(2) Similar. □

With this we can prove the following result, which is originally due to Barendregt and Koymans.

Theorem 9.

(1) *Every λ-algebra can be embedded in a λ model.*

(2) *Every λ-algebra is the homomorphic image of a λ model.*

Proof.

(1) $\mathcal{A} \cong \mathcal{M}^0(Th(\overline{\mathcal{A}})) \preccurlyeq \mathcal{M}(Th(\mathcal{A}))$.

(2) Consider $\varphi : \mathcal{M}(Th(\bar{\mathcal{A}})) \to \mathcal{M}^0(Th(\overline{\mathcal{A}})) \cong \mathcal{A}$, where φ is the onto map that replaces every free variable by $\mathbf{K}$. That this is a homomorphism follow directly from the definition of substitution. □

3. Category Theoretic Semantics

We will now give an alternative mathematical characterization of the λ models and λ–algebras. While this presentation uses the language of category theory rather than the language of algebra, and so is likely to be less familiar to those approaching this material from a background in the kind of logic typically taught in philosophy departments, it allows us to give a precise and elegant description of the relationship between the two sorts of λ-structures, and of the structures themselves.

We begin by showing how to define a syntactical applicative structure for any reflexive object in a cartesian closed category (ccc). Let $\mathbf{C}$ be a ccc with a reflexive object U via the maps $F : U \to U^U$ and $G : U^U \to U$. We will call our structure '$\mathcal{M}(\mathbf{C})$'. The domain of $\mathcal{M}(\mathbf{C})$ is $|U|$, the set of points of U. We define the map $\mathrm{Ap} : U \times U \to U$ to be the map $ev_{U,U} \circ (F \times 1_U)$. We use this to define the requisite binary operation on $|U|$ For $f, g : A \to U$, define $f \bullet_A g = \mathrm{Ap} \circ \langle f, g \rangle$. In the particular case of points $x, y \in |U|$, we put

$$x \bullet y = x \bullet_1 y = \mathrm{Ap} \circ \langle x, y \rangle.$$

Our applicative structure is $\langle |U|, \bullet \rangle$.

We want now to define the notion of an interpretation for our applicative structure. In §§1 and 2 of this chapter we have defined interpretations in terms of arbitrary valuations, where a valuation was a map taking the whole, infinite set of variables to the domain of the interpretation. While this was the natural approach to take in these cases, it is not natural on the category theoretical approach, where our valuations depend on the existence of appropriate products, and we are only guaranteed the existence of finite products in a ccc. Since we eventually want to be able to show that we are dealing with λ–algebras in the original sense, we will have to be able to handle arbitrary valuations. We will do this by making a detour through a definition of interpretation for sufficiently long sequences of variables. To this end we will proceed as follows.

Suppose $\mathbf{FV}(M) \subseteq \{\Delta\}$. (The reader is reminded of the various sorts of notation introduced in chapter 2 for sequences of variables.) Write 'T_{U^Δ}' for the unique map from U^Δ to $\mathbf{1}$. We define $\|M\|_\Delta : U^\Delta \to U$ inductively as follows.

$$\|x\|_\Delta = \pi_x^\Delta \qquad\qquad \text{for } x \in \mathbf{Var}$$

$$\|\mathbf{a}\|_\Delta = a \circ T_{U^\Delta} \qquad\qquad \text{for } a \in |U|$$

$$\|P(Q)\|_\Delta = \|P\|_\Delta \bullet_{U^\Delta} \|Q\|_\Delta$$

$$\|(\lambda x.P)\|_\Delta = G \circ \overline{\|P\|_{\Delta,x}}$$

where we assume by the variable convention, in the last clause, that $x \notin \Delta$.

Now to extend this so that it can handle arbitrary valuations $\varrho : \mathbf{Var} \to |U|$, let $\varrho^{\Delta} = \langle \varrho(x_1), \ldots, \varrho(x_n) \rangle$. We now define, for $\Delta = \mathbf{FV}(M)$,

$$\|M\|^{\varrho} = \|M\|_{\Delta} \circ \varrho^{\Delta}.$$

It is easily checked that $\|M\|^{\varrho} \in |U|$. Finally, we can define $\mathcal{M}(\mathbf{C})$ to be the structure $\langle |U|, \bullet, \|\cdot\| \rangle$.

However, having our structure in hand and showing that it is the structure we want are two very different matters. We must begin with some bookkeeping. First we check to be sure that our interpretation does not depend on choice of Δ.

Lemma 6. *Let* $\mathbf{FV}(M) \subseteq \{\Gamma\} \subseteq \{\Delta\}$. *Then* $\|M\|_{\Delta} = \|M\|_{\Gamma} \circ \Pi^{\Delta}_{\Gamma}$.

Proof. By induction on the structure of M. The interesting case is where $M \equiv (\lambda y.P)$, where we can argue:

$$\begin{aligned}
\|(\lambda y.P)\|_{\Gamma} \circ \Pi^{\Delta}_{\Gamma} &= G \circ \overline{\|P\|_{\Gamma,y}} \circ \Pi^{\Delta}_{\Gamma} \\
&= G \circ \overline{\|P\|_{\Gamma,y} \circ \left(\Pi^{\Delta}_{\Gamma} \times 1_U\right)} && \text{by Proposition 2.5(1)} \\
&= G \circ \overline{\|P\|_{\Gamma,y} \circ \left(\Pi^{\Delta,y}_{\Gamma,y}\right)} \\
&= G \circ \overline{\|P\|_{\Delta,y}} && \text{by IH} \\
&= \|(\lambda y.P)\|_{\Delta} && \square
\end{aligned}$$

As a stepping stone to proving Lemma 8 below we need

Lemma 7. *Take* $\vec{x} = \langle x_1, \ldots, x_n \rangle$ *and* $\vec{N} = \langle N_1, \ldots, N_n \rangle$, *and let* $\mathbf{FV}(M) \subseteq$

$\{\vec{x}\} = \{\Delta\}$ *and* $\mathbf{FV}(\vec{N}) \subseteq \{\Gamma\}$. *Then*

$$\|M[\vec{x}/\vec{N}]\|_\Gamma = \|M\|_\Delta \circ \langle \|N_1\|_\Gamma, \ldots, \|N_n\|_\Gamma \rangle .$$

Proof. The reader is reminded that we are assuming that the variable conventions are in place. We will write '$\left\langle \|\vec{N}\|_\Gamma \right\rangle$' for $\langle \|N_1\|_\Gamma, \ldots, \|N_n\|_\Gamma \rangle$. We will also need to appeal to the fact that

$$(\dagger) \qquad \left\langle \|\vec{N}\|_{\Gamma,y}, \|y\|_{\Gamma,y} \right\rangle = \left\langle \|\vec{N}\|_\Gamma \right\rangle \times 1_U .$$

This can be shown as follows. Starting with $\left\langle \|\vec{N}\|_{\Gamma,y}, \|y\|_{\Gamma,y} \right\rangle$, we can use Lemma 6 n-times to calculate that this is $\left\langle \left\langle \|N_1\|_{\Gamma,y} \circ \Pi_\Gamma^{\Gamma,y}, \ldots, \|N_n\|_{\Gamma,y} \circ \Pi_\Gamma^{\Gamma,y} \right\rangle \pi_y^{\Gamma,y} \right\rangle$. With another dose of n-fold perseverance we can use Proposition 2.5(2) to show that this is $\left\langle \|\vec{N}\|_\Gamma \circ \Pi_\Gamma^{\Gamma,y}, \pi_y^{\Gamma,y} \right\rangle$. This is easily seen to be $\left\langle \|\vec{N}\|_\Gamma \circ \pi_1, 1_U \times \pi_2 \right\rangle$, where π_1 and π_2 are the projections for $(U^\Gamma)^n \times U$. By the definition of $\times$ for arrows, this is $\left\langle \|\vec{N}\|_\Gamma \right\rangle \times 1_U$.

The main part of the proof is by an induction which is routine except for the case that goes like this:

$$\begin{aligned}
\|(\lambda y.P)[\vec{x}/\vec{N}]\|_\Gamma &= \|(\lambda y.P)[\langle \vec{x}, y\rangle / \langle \vec{N}, y\rangle]\|_\Gamma \\
&= G \circ \overline{\left(\|P[\langle \vec{x}, y\rangle / \langle \vec{N}, y\rangle]\|_{\Gamma,y}\right)} \\
&= G \circ \overline{\left(\|P\|_{\Delta,y} \circ \left\langle \|\vec{N}\|_{\Gamma,y}, \|y\|_{\Gamma,y} \right\rangle\right)} && \text{by IH} \\
&= G \circ \overline{\left(\|P\|_{\Delta,y} \circ \left\langle \|\vec{N}\|_\Gamma \right\rangle \times 1_U\right)} && \text{by } (\dagger) \\
&= G \circ \overline{(\|P\|_{\Delta,y})} \circ \left\langle \|\vec{N}\|_\Gamma \right\rangle && \text{by 2.5(1)} \\
&= \|(\lambda y.P)\|_\Delta \circ \left\langle \|\vec{N}\|_\Gamma \right\rangle . && \square
\end{aligned}$$

It is now straightforward to show

Lemma 8. *Let* $\mathbf{FV}((\lambda x.M)) \subseteq \{\Delta\}$, $\mathbf{FV}((\lambda x.M)(N)) \subseteq \{\Gamma\}$, *and* $\{\Delta\} \subseteq \{\Gamma\}$. *Then*

$$\|M[x/N]\|_\Gamma = \|M\|_{\Delta,x} \circ \left\langle \Pi^\Gamma_\Delta, \|N\|_\Gamma \right\rangle.$$

Proof. Apply lemma 7 with $\vec{x} = \langle \Delta, x \rangle$ and $M[\vec{x}/\vec{N}] \equiv M[\langle \Delta, x\rangle \,/\, \langle \Delta, N\rangle]$. □

We can now prove

Proposition 4. *Let* $M, N \in \Lambda(\mathcal{M}(\mathbf{C}))$ *and* $\mathbf{FV}(M) \cup \mathbf{FV}(N) \subseteq \{\Delta\}$. *Then*

$$\vdash M = N \implies \|M\|_\Delta = \|N\|_\Delta.$$

Proof. By induction on the length of the proof of $M = N$. The interesting steps are

(1) The (β) axiom. We can reason as follows:

$$\begin{aligned}
\|(\lambda x.P)(Q)\|_\Delta &= \left(G \circ \overline{\|P\|_{\Delta,x}}\right) \bullet_{U^\Delta} \|Q\|_\Delta \\
&= \mathrm{Ap} \circ \left\langle G \circ \overline{\|P\|_{\Delta,x}}, \|Q\|_\Delta \right\rangle \\
&= ev_{U,U} \circ (F \times 1_U) \circ \left\langle G \circ \overline{\|P\|_{\Delta,x}}, \|Q\|_\Delta \right\rangle \\
&= ev_{U,U} \circ \left\langle F \circ G \circ \overline{\|P\|_{\Delta,x}}, 1_U \circ \|Q\|_\Delta \right\rangle && \text{by 2.5(3)} \\
&= ev_{U,U} \circ \left\langle \overline{\|P\|_{\Delta,x}}, 1_U \circ \|Q\|_\Delta \right\rangle \\
&= ev_{U,U} \circ \overline{\|P\|_{\Delta,x}} \times 1_U \circ \langle 1_U, \|Q\|_\Delta \rangle && \text{by 2.5(3)} \\
&= \|P\|_{\Delta,x} \circ \langle 1_U, \|Q\|_\Delta \rangle \\
&= \|P[x/Q]\|_\Delta && \text{by lemma 8.}
\end{aligned}$$

(2) the rule (ξ). From $P = Q$ we get $\|P\|_\Delta = \|Q\|_\Delta$ by IH, so we can argue

$$\begin{aligned}
\|P\|_\Delta = \|Q\|_\Delta &\implies \|P\|_\Delta \circ \Pi_\Delta^{\Delta,x} = \|Q\|_\Delta \circ \Pi_\Delta^{\Delta,x} \\
&\implies \|P\|_{\Delta,x} = \|Q\|_{\Delta,x} \qquad \text{by lemma 6} \\
&\implies \overline{\|P\|_{\Delta,x}} = \overline{\|Q\|_{\Delta,x}} \\
&\implies \|(\lambda x.P)\|_\Delta = \|(\lambda x.Q)\|_\Delta. \qquad \square
\end{aligned}$$

Immediately from this proposition and the definition of $\|\cdot\|^\varrho$ we get

Theorem 10. *Every ccc* **C** *with a reflexive object U determines a λ-algebra $\mathcal{M}(\mathbf{C}) = \langle |U|, \bullet, \|\cdot\| \rangle$.* $\square$

Since we are now dealing with a λ-algebra, it follows by our results in section B that we are dealing with a combinatory algebra, and so we can choose **K** and **S**, and define **1** and **I** as usual.

We can now state and prove a theorem which sets out precisely the relationships between λ-algebras, λ-models, and *extensional* λ-models.

Theorem 11. *Let $\mathcal{M} = \mathcal{M}(\mathbf{C})$ be a λ-algebra as defined above with U a reflexive object in the ccc* **C** *via the maps F and G.*

(1) *Let $\mathbf{FV}(M) \subseteq \|\Delta\|$. Then $\|\mathbf{1}(M)\|_\Delta = G \circ F \circ \|M\|_\Delta$.*

(2) *U has enough points $\iff$ $\mathcal{M}$ is a λ-model.*

(3) *$U \cong U^U$ via F, G $\iff$ $\mathcal{M} \models \mathbf{1} = \mathbf{I}$.*

(4) *$U \cong U^U$ via F, G and U has enough points $\iff$ $\mathcal{M}$ is extensional.*

Proof.

For (1), we have

$$\begin{aligned}
\|\mathbf{1}(M)\|_\Delta &= \|(\lambda y.M(y))\|_\Delta && \text{by the remarks preceding Theorem 7}\\
&= G \circ \overline{\|M\|_{\Delta,y} \bullet_{U^{\Delta,y}} \|y\|_{\Delta,y}}\\
&= G \circ \overline{\mathrm{Ap} \circ \langle \|M\|_{\Delta,y}, \|y\|_{\Delta,y}\rangle}\\
&= G \circ \overline{ev_{U,U} \circ \langle F \circ \|M\|_{\Delta,y} \circ \|y\|_{\Delta,y}\rangle} && \text{by 2.5(3)}\\
&= G \circ \overline{ev_{U,U} \circ \left\langle F \circ \|M\|_\Delta \circ \Pi_\Delta^{\Delta,y}, \pi_\Delta^{\Delta,y}\right\rangle} && \text{by Lemma 6}\\
&= G \circ \overline{ev_{U,U} \circ \langle F \circ \|M\|_\Delta \circ \pi_1, \pi_2\rangle}
\end{aligned}$$

where π_1 and π_2 are the projections for $U^{\Delta,y}$. But by definition of $\times$ for arrows, this is $G \circ \overline{ev_{U,U} \circ (F \circ \|M\|_\Delta) \times 1_U}$. But $F \circ \|M\|_\Delta : U^U \to U^U$, and so is the transpose of a map $U^U \times U \to U$, which obviously must be $ev_{U,U} \circ (F \circ \|M\|_\Delta) \times 1_U$.

For (2), first suppose U has enough points. In Chapter 2 we mentioned that for any A, $\mathbf{1} \times A = A^1 \cong A$. With this in mind it is easy to show that if U has enough points then so too does $\mathbf{1} \times U$. Now let $a, b \in |U|$, and suppose that for any $x \in |M|$, $a \bullet x = b \bullet x$, i. e. $\mathrm{Ap} \circ \langle a, x\rangle = \mathrm{Ap} \circ \langle b, x\rangle$. By the definition of Ap and 2.5(3),

$$ev_{U,U} \circ \langle F \circ a, x\rangle = ev_{U,U} \circ \langle F \circ b, x\rangle ,$$

so by 2.5(3) again, $ev_{U,U} \circ (F \circ a) \times 1_U \circ \langle 1_1, x\rangle = ev_{U,U} \circ (F \circ b) \times 1_U \circ \langle 1_1, x\rangle$. But there is a bijective correspondence between points x of U and maps $\langle 1_1, x\rangle : \mathbf{1} \times \mathbf{1} \to \mathbf{1} \times U$ and so, by the isomorphism $\mathbf{1} \times \mathbf{1} \cong \mathbf{1}$, to the points of $\mathbf{1} \times U$. Since $\mathbf{1} \times U$ has enough points,

$$ev_{U,U} \circ (F \circ a) \times 1_U = ev_{U,U} \circ (F \circ b) \times 1_U ,$$

and so

$$\overline{ev_{U,U} \circ (F \circ a) \times 1_U} = \overline{ev_{U,U} \circ (F \circ b) \times 1_U}.$$

But in the proof of (1) we saw that, for any point c, $\mathbf{1} \bullet c = G \circ F \circ c = G \circ \overline{ev \circ (F \circ c) \times 1_U}$, hence $\mathbf{1} \bullet a = \mathbf{1} \bullet b$ and $\mathcal{M}$ is a λ-model.

For ($\Longleftarrow$), suppose $\mathcal{M}$ is a λ-model, $f, k : U \to U$ and that for any $x \in |U|$, $f \circ x = k \circ x$. This gives us two arrows $f \circ \pi_2, k \circ \pi_2 : U \times U \to U$, from which we get two more arrows $\overline{f \circ \pi_2}, \overline{k \circ \pi_2} : U \to U^U$ such that, for each x, $G \circ \overline{f \circ \pi_2} \circ x = G \circ \overline{k \circ \pi_2} \circ x$. It is not hard to calculate that, for each x, $G \circ \overline{f \circ \pi_2} \bullet x = G \circ \overline{f \circ \pi_2} \circ x$, and similarly for k in place of f. By the first half of this proof we have

$$\mathbf{1} \bullet G \circ \overline{f \circ \pi_2} = \mathbf{1} \bullet G \circ \overline{k \circ \pi_2}.$$

This yields

$$G \circ F \circ G \circ \overline{f \circ \pi_2} = G \circ F \circ G \circ \overline{k \circ \pi_2}$$

by (1). Applying F to both sides leaves us with an equality, and since $F \circ G = 1_{U^U}$, $\overline{f \circ \pi_2} = \overline{k \circ \pi_2}$ from which we have $f \circ \pi_2 = g \circ \pi_2$, whence $f = g$. Therefore U has enough points.

Next we show the ($\Longrightarrow$) half of (3). By (1) we have $\|\mathbf{1}(M)\|_\Delta = \|M\|_\Delta$, since $G \circ F = 1_U$. In particular, $\|\mathbf{1}(x)\|_\Delta = \|x\|_\Delta$. So again appealing to the remarks before Theorem 7 we can argue that

$$\|\mathbf{1}\|_\Delta = \|(\lambda x.\mathbf{1}(x))\|_\Delta = \|(\lambda x.x)\|_\Delta = \|\mathbf{I}\|_\Delta,$$

i.e. $\mathcal{M} \models \mathbf{1} = \mathbf{I}$.

Now ($\Longleftarrow$). Assume $\mathcal{M} \models \mathbf{1} = \mathbf{I}$. Then it is straightforward to calculate that

$\|\mathbf{1}\| = \|(\lambda x.(\lambda y.x(y)))\|$, which by (1) is $G \circ \overline{\|(\lambda y.x(y))\|_x}$. By the argument for (1), this is $G \circ \overline{G \circ F \circ \|x\|_x} = G \circ \overline{G \circ F \circ \pi_x^x}$. Similarly, $\|\mathbf{I}\| = G \circ \overline{\pi_x^x}$. So $G \circ \overline{G \circ F \circ \pi_x^x} = G \circ \overline{\pi_x^x}$. But as we saw in the first half of this proof, G is monic, so $\overline{G \circ F \circ \pi_x^x} = \overline{\pi_x^x}$ hence $G \circ F \circ \pi_x^x = \pi_x^x$. Now $\pi_x^x = \pi_2 : 1 \times U \to U$ so, for $g : V \to U$, $g = \pi_x^x \circ \langle !, g \rangle$, so $G \circ F \circ g = g$. On the other hand, for $f : U \to V$, $f \circ G \circ F = f \circ G \circ F \circ \pi_x^x \circ \langle !, 1_U \rangle = f \circ \pi_x^x \circ \langle !, 1_U \rangle = f$. Thus $G \circ F = 1_U$.

Finally, we get (4) by (2) and (3) together with Proposition 2. □

So far we have seen that every *ccc* with a reflexive object gives rise to a λ-algebra, and if the reflexive object has enough points, then it gives rise to a λ-model. Conversely, every λ-algebra can be obtained from a reflexive object in a *ccc*, and every λ-model from one with enough points. We will satisfy ourselves here with a quick sketch of how this is proved.

Let $\mathcal{A} = \langle X, \mathbf{R}, \| \cdot \| \rangle$ be a λ–algebra. The ***Karoubi envelope*** of $\mathcal{A}$, which we denote by '$\mathbf{C}(\mathcal{A})$)', is the category defined by:

(1) For each $a, b \in \mathcal{A}$, $a \circ b = \|(\lambda x.a)(b(x))\|$

(2) $\mathrm{Ob}(\mathbf{C}(\mathcal{A})) = \{\, a \in \mathcal{A} \mid a \circ a = a \,\}$

(3) $Hom(a, b) = \{\, f \in \mathcal{A} \mid b \circ f \circ a = f \,\}$

(4) For each $a \in \mathcal{A}$, $1_a = a$

(5) The composition of $f : a \to b$ and $g : b \to c$ is $g \circ f$.

This is easily shown to be a category. Dana Scott in [Scott 1980b] showed the fairly simple method (once someone had thought of it) for defining products, exponents and a terminal object, that $\mathbf{C}(\mathcal{A})$ is a *ccc*, that the objects $\mathbf{1}$ and $\mathbf{I}$ are easily defined, and that **I** is a reflexive object via $F = G = \mathbf{1}$. Koymans showed shortly thereafter that where $\mathcal{M}$ is the algebra defined as in this section for $\mathbf{C}(\mathcal{A})$ with reflexive object **I** via **1** and **1**, $\mathcal{M} \cong \mathcal{A}$. See [Koymans] or [Barendregt 1984, pp. 114–15] for a proof.

4. Example of a λ-Model

Consider $\mathcal{P}(\omega) = \{\, x \mid x \subseteq \mathbb{N} \,\}$, the power set of the natural numbers ordered by $\subseteq$. We discuss this set with its *Scott topology*, for which

$$O_{e_n} = \{\, x \mid e_n \subseteq x \,\}$$

forms a base, where $\langle e_0, e_1, \dots \rangle$ is some arbitrary but fixed enumeration of the *finite* subsets of $\mathbb{N}$. That is to say the *open sets* of the Scott topology are all the unions of these basic open sets. We will circumvent the problems discussed at the start of this chapter by considering only the continuous functions with respect to this topology, rather than the whole set theoretic function space, so that we can define a pair of maps to show that $\mathcal{P}(\omega)$ is a reflexive object. It is not hard to show that

Proposition 5. *For* $f : \mathcal{P}(\omega) \to \mathcal{P}(\omega)$

(1) f *is continuous* $\implies f$ *is monotonic.*

(2) f *is continuous* $\iff f(x) = \bigcup \{\, f(e) \mid e \subseteq x, e$ *finite* $\}$

$\iff$ *for all* $x \in \mathcal{P}(\omega)$ *and any* e_m, $e_m \subseteq f(x)$ *iff* $\exists\, e_n \subseteq x \,[e_m \subseteq f(e_n)]$. □

Now, we can easily code the ordered pairs and finite subsets of $\mathbb{N}$. Indeed, we assumed such ability when we spoke above of an enumeration of the finite subsets.

By (2) f is determined by its values on finite sets. Thus we can get a map graph : $[\mathcal{P}(\omega) \to \mathcal{P}(\omega)] \to \mathcal{P}(\omega)$ with an obvious inverse. So we define, for $f \in [\mathcal{P}(\omega) \to \mathcal{P}(\omega)]$,

$$\text{graph}(f) = \{ \langle n, m \rangle \mid m \in f(e_n) \},$$

and for $u \in \mathcal{P}(\omega)$ we define $\text{fun}(u) \in [\mathcal{P}(\omega) \to \mathcal{P}(\omega)]$ by

$$\text{fun}(u)(x) = \{ m \mid \exists e_n \subseteq x[\langle n, m \rangle \in u] \}.$$

It is not hard to prove

Proposition 6.

(1) graph : $[\mathcal{P}(\omega) \to \mathcal{P}(\omega)] \to \mathcal{P}(\omega)$ *is continuous.*

(2) *For all* $u \in \mathcal{P}(\omega)$, $\text{fun}(u) \in [\mathcal{P}(\omega) \to \mathcal{P}(\omega)]$.

(3) fun : $\mathcal{P}(\omega) \to [\mathcal{P}(\omega) \to \mathcal{P}(\omega)]$ *is continuous.*

(4) *For any* $f \in [\mathcal{P}(\omega) \to \mathcal{P}(\omega)]$, $\text{fun}(\text{graph}(f)) = f$.

(5) *For any* $u \in \mathcal{P}(\omega)$, $\text{graph}(\text{fun}(u)) = u$. □

By the earlier results, then, $\mathcal{P}(\omega)$ is a λ-model, since putting F = fun and G = graph shows that $\mathcal{P}(\omega)$ is a reflexive object via G and F and $F \circ G = 1_{\mathcal{P}(\omega)}$.

5. Philosophical Interlude

We are now in a position to see the point of some of the disagreements over whether λ-algebras or λ-models are a more appropriate structures to serve as the basic tools for interpreting the λ-calculus. The corollary to Theorem 4 and the comments at the end of §1 are the source of most of the complaints against λ-algebras. For example, Hindley and Seldin point out that "we can make deductions by rule (ξ), for example

$$(\lambda x.y(x)) = y \vdash (\lambda y.(\lambda x.y(x))) = (\lambda y.y).$$

So any reasonable definition of 'model of [λ-calculus]' should have the property that if $\mathcal{D}$ is such a model, then

$$(\mathcal{D} \models (\lambda x.y(x)) = y) \implies (\mathcal{D} \models (\lambda y.(\lambda x.y(x))) = (\lambda y.y).$$

But this property fails for [λ-algebras]." [Hindley and Seldin, p.111].

The other main argument levelled against the λ-algebras is that all of the most natural mathematical models of the λ-calculus (including $\mathcal{P}(\omega)$) are λ-models, and not merely λ-algebras.

This has by no means convinced all those who work on the λ-calculus that they ought to restrict their attention to the λ-models. The main arguments in favour of the λ-algebras are historical. The λ-calculus and its close cousin combinatory logic were originally invented by mathematicians who were consciously and explicitly working with a notion of *function* or *operation* different from the set theoretic one which has come to dominate twentieth century mathematics. For instance, Church says, in a passage Dana Scott likes to quote,

> Here, however, we regard the operation or rule of correspondence, which constitutes the function, as being first given, and the range of arguments then determined as consisting of the things to which the operation is applicable. This is a departure which is natural in passing from consideration of functions in a special domain to the consideration of functions in general, and it finds its support in consistency theorems [i.e.

> the Church–Rosser Theorems] which will be proved below. ([Church], quoted in [Scott 1980a, p.231])

Now, as Scott points out, there are considerable problems involved in taking the λ-calculus as we have presented it here (rather than the more restrictive λ-I-calculus [Church] originally worked with) as fitting this philosophical stance very well. The λ-calculus we have presented, which is the calculus normally worked with nowadays, requires every operation to be applicable to every other, and so the range of application is indeed fixed in advance. Furthermore, it's not very clear why Church thinks that the Church–Rosser Theorem provides a justification for the claim that any version of the λ-calculus provides an account of functions in general. At best it shows that the formalization meets one necessary condition on an adequate account.

Nonetheless, this quote is of interest to us because it makes clear that Church, Curry and the other founders of the λ-calculus had some other notion of function in mind than the set-theoretic one. And whatever we think of their attempts to clarify the notion of "function as rule", it is clear that the appropriate notion will be *intensional*.

The point of this historical digression is this. If we intend to take the notion of function *intensionally*, then we ought not to be surprised that the model theory is less smooth in some respects than if we take it extensionally. This is simply what happens if one works with the notion of *rule* instead of the notion of *graph*. By keeping both the notions of λ-algebra and of λ-model at out disposal, we can treat weak extensionality as a simplifying assumption, and by contrasting the λ-algebras and the λ-models we can get some idea of exactly what this simplifying assumption buys for us.

Some of the explicit defenses of λ-algebras in the literature sound a bit tendentious. For instance, Koymans says "[Weak extensionality] seems to be a reasonable assumption for a general lambda calculus model, were it not that the structure of all closed lambda terms (which certainly should be a model) does not satisfy the above axiom" [Koymans, p.3]. However, a more sympathetic reading of Koymans' parenthetical remark can be had if we are willing to give it a historical spin. After all, as was mentioned at the start of this chapter, the term models, open and closed, were effectively the only models for the λ-calculus for at least thirty years. So we can read Koymans as refusing to convict those who worked with the closed term model for all those years of being confused.

Perhaps the most appropriate way to respond to the challenge raised by Hindley and Seldin is to respond in kind. The λ-algebras also possess some nice model theoretic properties not shared by the λ-models. For instance we mention the following, which follows easily from the fact that the class of λ-algebras is equationally definable (see p. 94 of [Barendregt 1984]) and the fact that the closed term model is a λ-algebra but not a λ-model.

Proposition 7.

(1) *If $\mathcal{M}$ is a combinatory algebra or λ-algebra, then so is $\mathcal{M}^0$.*

(2) *There exists a λ-model $\mathcal{M}$ such that $\mathcal{M}^0$ is not a λ-model.* □

It follows that we have a Submodel Theorem for λ–algebras but not for λ–models. That is, if $\langle X, \circ, s, k\rangle$ is a λ–algebra, then for $X' \subseteq X$ with $s, k \in X'$, the substructure $\langle X', \circ, s, k\rangle$ is a λ–algebra. No similar claim can be made for λ–models. That is, the λ–algebras form a *variety* in the sense of universal algebra, and so they ought not be dismissed as mathematically uninteresting as a class.

The fact that the λ–models do not form a variety reflects the fact, already mentioned, that in the λ–models the abstraction operator has the internal strength of a universal quantifier.

CHAPTER V

FIRST ORDER LOGIC: SYNTAX AND SYSTEMS

In this chapter we will set out the syntax for a first order language with a single variable binding term forming operator. We will use 'σ' as an arbitrary operator. Whether σ is ε, or τ, or some other operator will be determined by the axioms we add which dictate the operator's behaviour. At the end of the chapter we will set out the rules for most of the logical systems that will occupy us in Chapters 6–11.

For the most part, the definitions in this chapter are close kin of those to be found in any standard treatment of first order logic, for instance the one contained in [Bell and Machover]. There is actually quite a bit of finicky work that needs to be done to ensure that the syntactical machinery works, but this probably doesn't make for exciting reading for most people. So a few words about what kinds of modifications need to be made to the standard definitions are probably in order so that those inclined to just skim this section will know what they're missing.

First of all, it will be convenient to use the same language for both the intuitionistic and classical cases. We will therefore take all the connectives '$\vee$', '$\wedge$', '$\Rightarrow$', '$\neg$', and both the quantifiers '$\forall$' and '$\exists$' as primitive symbols of the language[1]. Since the extra length this engenders for metatheoretical proofs in the classical case will

[1]Note that we use the symbol '$\Rightarrow$' as the symbol for implication in $\mathcal{L}$, while the long arrow '$\Longrightarrow$' will be a metalinguistic symbol for implication.

for the most part be handled by hand-waving, the result will be a more economical presentation. Similarly, we will take the quantifiers and σ all to be primitive. It is well-known that in the (extensional) classical case one can define the quantifiers in terms of ε. However, we cannot define the quantifiers in terms of ε in the intuitionistic case, as we shall see. We will also see in Chapter 8 that the systems which result by taking the quantifiers to be primitive and defined, respectively, in non-extensional classical cases are not precisely equivalent. So defining the quantifiers in terms of ε in one case would detract significantly from the uniformity of this presentation without any particular benefit.

We will work almost exclusively with a language without function symbols, but with equality. As a result, the only terms will be variables, constants, and σ-terms. In contexts where we need function symbols we will explicitly add them.

As has already been mentioned, σ is a variable binding operator which makes terms out of formulas. One result of this is that the definitions of terms and formulas cannot be neatly compartmentalized in the way they are in standard treatments of first order logic. So many of the bookkeeping definitions will require one elaborate inductive definition rather than two definitions, one for each case. Similarly, we will need to define the notion of a term being *free for a variable* for occurrences of variables within terms, not just in formulas. The result is that the sections on substitution and variable binding take some work.

Readers familiar with the approach to syntax taken in [Leisenring] should note that while this approach is simpler than ours and is well suited to Leisenring's purposes, it won't do for us. The problem is that Leisenring's approach depends on certain extensionality assumptions being taken as axioms of the ε-calculus. This allows him to avoid the need for relabelling bound variables in certain contexts. Since we wish to avoid these extensionality assumptions, we need to grind out the details as we do below.

It is also probably worth noting that the account below is more complicated than the usual accounts of, for instance, definite descriptions (cf., e. g. [Bencivenga et al., pp. 215–23]) because we allow *open* σ-terms, and not just closed ones. This is very important, since several of the proofs below depend on the availability of these open terms. The syntax would be much simpler if we were to restrict ourselves to closed terms.

Finally, the reader should note that when giving inductive definitions we will be assuming that the defined sets are the smallest ones meeting the specified conditions, sometimes without explicitly mentioning it.

1. The Language $\mathcal{L}$

We begin by describing an arbitrary *first order language $\mathcal{L}$ for the σ-calculus*.

The symbols of $\mathcal{L}$ include all and only the members of the following disjoint sets.

(1) **Var** $= \{v_1, v_2, v_3, \dots\}$, a countably infinite sequence of (*individual*) *variables*.

(2) **Con** $= \{c_1, c_2, c_3, \dots\}$, a set, which might be infinite, finite or empty of (*individual*) *constants*.

(3) **Pred** $= \bigcup_{1 \leqslant n < \omega} \mathbf{Pred}_n$ a set of predicate symbols, where for each positive

integer n $\mathbf{Pred}_n$ is a possibly empty set of n *ary predicate symbols*.

(4) The *connectives* $\neg$ (negation), $\Rightarrow$ (implication), $\vee$ (disjunction), and $\wedge$ (conjunction).

(5) The *universal quantifier* $\forall$ and the *existential quantifier* $\exists$.

(6) The *identity predicate* $=$.

(7) The *σ-operator* σ.

(8) The punctuation symbols ., (, and).

We count the connectives, quantifiers, identity predicate, and σ as the *logical symbols* of the σ-calculus, and these are assumed to be the same in all first order σ-languages. The predicate symbols and constants are *extralogical symbols*, and they vary from one language to another.

We will adopt the following *notational conventions*. We will use 'x', 'y', 'z', 'u', 'v', and 'w', possibly with subscripts, to denote arbitrary variables. The letters 'a', 'b', and 'c', possibly with subscripts, will be used to denote arbitrary constants. We denote syntactical identity by '$\equiv$', so $M \equiv N$ means that M and N are exactly the same expression.

An *expression of length n* is any string $p_1, \ldots, p_n$ of symbols of $\mathcal{L}$. If $n = 0$, we denote the expression by e. If S and T are expressions, ST is the expression formed by concatenating S and T. If $R \equiv ST$, then S is an *initial segment* and T is a *terminal segment* of R. If $S \not\equiv e$, T is a *proper terminal segment* of R. If $T \not\equiv e$, S is a *proper initial segment* of R.

In most treatments of first order logic it is standard procedure to define the terms and formulas by two separate recursive definitions. Because we now must take into account terms formed out of formulas, it is better to define both sets at once.

Definition 1. The *well-formed expressions* of $\mathcal{L}$ fall into two disjoint sets, the set of *well-formed terms* (**Terms**), and the set of *well-formed formulas* (**Wffs**), which are the smallest sets such that:

(1) $\mathbf{Var} \subseteq \mathbf{Terms}$

(2) $\mathbf{Con} \subseteq \mathbf{Terms}$

(3) $t_1, \ldots t_n, \in \mathbf{Terms}, \quad P \in \mathbf{Pred}_n \implies Pt_1 \ldots t_n \in \mathbf{Wffs}$

(4) $s, t \in \mathbf{Terms} \implies (s = t) \in \mathbf{Wffs}$

(5) $\varphi, \psi \in \mathbf{Wffs} \implies \neg\psi, (\varphi \vee \psi), (\varphi \wedge \psi), (\varphi \Rightarrow \psi) \in \mathbf{Wffs}$

(6) $x \in \mathbf{Var}, \varphi \in \mathbf{Wffs} \implies \exists x.\varphi, \forall x.\varphi \in \mathbf{Wffs}$

(7) $x \in \mathbf{Var}, \varphi \in \mathbf{Wffs} \implies \sigma x.\varphi \in \mathbf{Terms}$

Expressions which qualify as well-formed by virtue of (7) are called *σ-terms*. Those which qualify by (6) are called *quantified formulas*, and the formula $\forall x.\varphi$ (or $\exists x.\varphi$) is called the *scope* of that occurrence of $\forall x$ ($\exists x$). Formulas which qualify by (3) and (4) are called *atomic formulas*. A formula is *σ-free* if the symbol 'σ' does not occur in it, *identity free* if '=' does not occur in it, and *elementary* (or *atomic*) if none of '$\forall$', '$\exists$', and 'σ' occurs in it.

We will use 's', 't', and 'r', possibly with subscripts or primes, to denote arbitrary terms of $\mathcal{L}$, and lower case Greek letters, usually 'φ', 'ψ', 'α', 'β' and 'γ', sometimes with subscripts or primes, to denote arbitrary well-formed formulas of $\mathcal{L}$. We use 'M' and 'N', possibly subscripted or primed, for arbitrary well-formed expressions.

We will use the following abbreviation:

$$(\varphi \Leftrightarrow \psi) \qquad \text{for} \qquad ((\varphi \Rightarrow \psi) \wedge (\psi \Rightarrow \varphi)).$$

We will often replace parentheses by brackets of various sizes to facilitate reading. Parentheses will often be omitted in accord with the author's peculiar notion of readability, subject to the convention that '$\Leftrightarrow$', '$\Rightarrow$', '$\vee$', and '$\wedge$' should be read in this order of priority, and that the scopes of '$\neg$', '$\forall$', '$\exists$', and 'σ' should be as short as possible. For formulas of the form $\varphi_1 \vee \cdots \vee \varphi_n$, $\varphi_1 \wedge \cdots \wedge \varphi_n$, or $\varphi_1 \Rightarrow \ldots \Rightarrow \varphi_n$ we will adopt the rule of ***association to the right*** (that is, we give the highest priority to the leftmost of a string of similar binary connectives). So, for example,

$$\varphi \wedge \psi \Rightarrow \beta \Leftrightarrow \neg\varphi \Rightarrow \psi \vee \beta \quad \text{is} \quad ([(\varphi \wedge \psi) \Rightarrow \beta] \Leftrightarrow [\neg\psi \Rightarrow (\psi \vee \beta)])$$

$$\forall x.\varphi \vee \beta \wedge \psi \quad \text{is} \quad (\forall x.\varphi \vee (\beta \wedge \psi))$$

$$(\varphi \Rightarrow \psi \Rightarrow \beta) \Rightarrow (\varphi \Rightarrow \psi) \Rightarrow \varphi \Rightarrow \psi \text{ is } ((\varphi \Rightarrow (\psi \Rightarrow \beta)) \Rightarrow ((\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow \psi)).$$

Note also that we cannot omit the parentheses from $(\varphi \wedge \beta) \wedge \psi$, $\forall x.(\varphi \wedge \psi)$, or $\varphi \wedge (\beta \vee \psi)$.

2. Free and Bound Variables, and Substitution

We began the corresponding section on the λ-calculus with a definition of a subterm. Things are not so simple here for a couple of reasons. The λ-calculus is a term calculus, which is to say that there is no distinction between formulas and terms. In first order logic there is no way to avoid this distinction, so we will need a definition of a subterm and of a subformula, as well as a definition of what it means for a term to occur in a formula. Also, while in standard treatments of first order logic we can give separate definitions for subterms and subformulas, the definitions depend on one another in the σ-calculus.

Definition 2. Let $s, t_1, \ldots, t_n \in$ **Terms**, $\varphi, \psi, \beta \in$ **Wffs**. We give a simultaneous recursive definition of the notions *s is a subterm of t*, *β is a subformula of φ*, and *s occurs in φ* as follows.

(1) *s is a subterm of s.*

(2) *β is a subformula of β.*

(3) $\varphi \equiv Pt_1, \ldots, t_n$ $(t_1 = t_2)$ is an atomic formula $\Longrightarrow$ $t_1, \ldots, t_n$ (t_1 and t_2) *occur in φ.*

(4) β is a subformlas of φ $\Longrightarrow$ *β is a subformula of $\neg\varphi$.*

(5) β is a subformula of φ or ψ $\Longrightarrow$ *β is a subformula of* $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$, and $(\varphi \Rightarrow \psi)$.

(6) t occurs in φ or in ψ $\Longrightarrow$ *t occurs in* $\neg\varphi$, $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$, and $(\varphi \Rightarrow \psi)$.

(7) β is a subformula of φ $\Longrightarrow$ *β is a subformula of* $\exists x.\varphi$ and $\forall x.\varphi$.

(8) t occurs in φ $\Longrightarrow$ *t occurs in* $\exists x.\varphi$ and $\forall x.\varphi$.

(9) t occurs in φ $\Longrightarrow$ *t is a subterm of $\sigma x.\varphi$.*

If s is a subterm of t and $s \not\equiv t$, we say s is a *proper subterm* of t. If β is a subformula of φ and $\beta \not\equiv \varphi$, then β is a *proper subformula* of φ.

For a well-formed expression M we define the *complexity of M* (and we write 'comp(M)') for the number of occurrences of $\neg$, $\vee$, $\wedge$, $\Rightarrow$, $=$, $\exists$, $\forall$ and σ in M.

In the name of bookkeeping, we prove the following lemma.

Lemma 1. *No proper initial segment of a well-formed expression is a well-formed expression.*

Proof. By induction on comp(M). Suppose comp(M) = 0. If $M \equiv x$ or $M \equiv c$, then there is no proper initial segment but ϵ, which is not a well-formed expression. If $M \equiv Pt_1 \dots t_n$ where $P \in \mathbf{Pred}_n$, each t_i for $i = 1, \dots, n$ is either a variable or a constant. So any proper initial segment (other than ϵ) doesn't have the appropriate number of terms.

Now suppose comp(M) = n. If the main operator in M is $\vee$, $\wedge$, $\Rightarrow$, or $=$ then any proper initial segment of M has more left parentheses than right parentheses, which clearly can't happen in a well-formed expression. If $M \equiv \neg\varphi$, $M \equiv \forall x.\varphi$, $M \equiv \exists x.\varphi$ or $M \equiv \sigma x.\varphi$, then $\varphi \in \mathbf{Wffs}$. Suppose $\varphi \equiv RT$ with $T \not\equiv \epsilon$ and $\neg R$, $\forall x.R$, $\exists x.R$ or $\sigma x.R$ is a well-formed expression. Then $R \in \mathbf{Wffs}$, which violates the inductive hypothesis. □

One useful consequence of this lemma is that it ensures that if we concatenate n terms to get $t_1 \dots t_n$, each t_i is uniquely determined in the sense that if we start peeling off well-formed terms from the left we will get precisely t_1 to t_n back again, and so clauses (3) of Definitions 2 and 3 are well behaved. This will also ensure that we can talk about variables that are members of the string without risk of confusion with occurrences of the same variable that are embedded, for instance, within a σ term. We rely on this fact, for example, in the first clause of the Definition 3 below.

We have still to deal with free and bound occurrences of variables, and with substitution. Ours will obviously need to be somewhat more complicated than the usual definitions, since binding now takes place in terms as well as in formulas. Also, we will need a notion of a free occurrence of an arbitrary term for some of our work below.

Definition 3. A given occurrence of a variable in a formula φ is a *free* occurrence if and only if it is not a *bound* occurrence in φ. A given occurrence of a variable in a σ-term t is a *free* occurrence in t if and only if it is not a *bound* occurrence in t.

(1) $\varphi \equiv Pt_1 \ldots t_n$ and $x = t_i$ for some $1 \leqslant i \leqslant n$ or $\varphi \equiv t_1 = t_2$ and $x \equiv t_1$ or $x \equiv t_2 \implies$ that occurrence of x is *free* in φ.

(2) $\varphi \equiv \neg\psi$ and a given occurrence of x is free in $\psi \implies$ that occurrence of x is free in φ.

(3) $\varphi \equiv \psi \vee \beta$, $\varphi \equiv \psi \wedge \beta$, or $\varphi \equiv \psi \Rightarrow \beta$ and a given occurrence of x is free in ψ or $\beta \implies$ that occurrence of x is free in φ.

(4) $\varphi \equiv \forall x.\beta$ or $\varphi \equiv \exists x.\beta \implies$ every occurrence of x is bound in φ.

(5) $t \equiv \sigma x.\beta \implies$ every occurrence of x in t is bound in t.

(6) $\varphi \equiv \forall y.\beta$ or $\varphi \equiv \exists y.\beta$ and $y \not\equiv x$ and a given occurrence of x is free in β $\implies$ that occurrence of x is *free* in φ.

(7) $t \equiv \sigma x.\beta$ and $y \not\equiv x$ and a given occurrence of x is free in $\beta \implies$ that occurrence of x is free in t.

(8) $\varphi \equiv Pt_1 \ldots t_n$ and an occurrence of x is free in t_i for some i, $1 \leqslant i \leqslant n$ $\implies$ that occurrence of x is free in φ.

The *set of free variables of* M, denoted by '$\mathbf{FV}(M)$', is the set of variables which have a free occurrence in M.

An occurrence of a term t in M is *a free occurrence of t in M* if there is no variable x such that $x \in \mathbf{FV}(t)$ and that instance of t occurs in some subformula or subterm N of M and $x \notin \mathbf{FV}(N)$. We define the *set of free terms of M*, denoted by '$\mathbf{FT}(M)$', to be the set of terms which have a free occurrence in M *which is not a proper subterm of a free occurrence of a term in M*. Note that this means that it will **not** in general be the case that $\mathbf{FV}(M) \subseteq \mathbf{FT}(M)$. This will give us some much needed flexibility in later chapters when discussing various interpretations of the σ–terms.

The next thing that we need is a definition of substitution. This takes a fair bit of work in texts for first order logic that don't merely wave their hands at the problem, and things of course only get more complicated for us with the increased opportunities for variable collision our extra bound variables give us. So we will have to spend some time on preliminaries.

We first specify what it is for a term to be free for a variable in a particular well-formed expression.

Definition 4. Let $s, t_1, \ldots, t_n \in$ **Terms**, $\varphi, \psi, \beta \in$ **Wffs**, and let t be an arbitrary term.

(1) $s \in$ **Var** or $s \in$ **Con** $\Longrightarrow$ *t is free for x in s.*

(2) $\varphi \equiv Pt_1 \ldots t_n$ and for $i = 1, \ldots, n$, t is free for x in t_i $\Longrightarrow$ *t is free for x in φ.*

(3) $\varphi \equiv t_1 = t_2$ and for $i = 1, 2$, t is free for x in t_i $\Longrightarrow$ *t is free for x in φ.*

(4) $\varphi \equiv \neg\psi$ and t is free for x in $\psi \implies$ *t is free for x in φ.*

(5) $\varphi \equiv \psi \vee \beta$, $\varphi \equiv \psi \wedge \beta$, or $\varphi \equiv \psi \Rightarrow \beta$ and t is free for x in ψ and in $\beta \implies$ *t is free for x in φ.*

(6) $\varphi \equiv \forall y.\beta$ or $\varphi \equiv \exists y.\beta$ and either (i) $x \notin \mathbf{FV}(\varphi)$ or (ii) $x \in \mathbf{FV}(\varphi)$ (and so $x \not\equiv y$) and y does not occur in t and t is free for x in $\beta \implies$ *t is free for x in φ.*

(7) $s \equiv \sigma y.\beta$ and either (i) $x \notin \mathbf{FV}(s)$ or (ii) $x \in \mathbf{FV}(s)$ (and so $x \not\equiv y$) and y does not occur in t and t is free for x in $\beta \implies$ *t is free for x in s.*

Note that it follows from (1)–(3) that if φ is atomic and σ–free, then t is free for x in φ, which allows the recursion to get off the ground.

Since we want to reserve the notation '$\varphi[x/t]$' for proper substitution, we will use the notation '$\varphi[[x/t]]$' ('$s[[x/t]]$') to denote the expression which results if every free occurrence of x in φ (in s) is replaced by an occurrence of t. More generally, we use '$\varphi[[t/t']]$' ('$s[[t/t']]$') to denote the expression which results if every free occurrence of t in φ (in s) is replaced by an occurrence of t'.

If z is not free in β, but is free for x in β then $\forall z.(\beta[[x/z]])$ is a *letter change* of $\forall x.\beta$. Under the same conditions $\exists z.(\beta[[x/z]])$ and $\sigma z.(\beta[[x/z]])$ are *letter changes* for $\exists x.\beta$ and $\sigma x.\beta$ respectively. Of course, this relationship is symmetric, since x must be free for z in $\beta[[x/z]]$.

If φ has a universal or existential subformula ψ, or if a σ–term s occurs in φ, then the replacement of an occurrence of ψ by a letter change of it, or of an occurrence of s by a letter change of it, is called a ***change of bound variables in*** *φ*. For a σ–term

$s \equiv \sigma x.\varphi$, each letter change for s is a *change of bound variables for* s; and if φ' results from φ by a change of bound variables, then $\sigma x.\varphi'$ is a *change of bound variables for* s.

These definitions are useful mainly because they allow us to define the notion of an alphabetic variant, which we turn to now. Note that in the first clause if φ is σ-free, then φ is its own only alphabetic variant.

Definition 5 (Alphabetic Variants).

(1) φ atomic and φ' results from φ by a finite sequence of changes of bound variables $\Longrightarrow$ φ' is an *alphabetic variant* of φ.

(2) $\varphi \equiv \neg\psi$ and ψ' is an alphabetic variant of ψ $\Longrightarrow$ $\neg\psi'$ is an *alphabetic variant* of φ.

(3) $\varphi \equiv \psi \vee \beta$, $\varphi \equiv \psi \wedge \beta$, or $\varphi \equiv \psi \Rightarrow \beta$ and ψ' and β' are alphabetic variants of ψ and β respectively $\Longrightarrow$ $\psi' \vee \beta'$, $\psi' \wedge \beta'$, or $\psi' \Rightarrow \beta'$ is an *alphabetic variant* of φ.

(4) $\varphi \equiv \forall x.\beta$ or $\varphi \equiv \exists x.\beta$ $\Longrightarrow$ (i) letter changes of φ are *alphabetic variants* of φ; and (ii) if β' is an alphabetic variant of β, then $\forall x.\beta'$ or $\exists x.\beta'$ is an *alphabetic variant* of φ.

(5) $s \equiv \sigma x.\beta$ $\Longrightarrow$ (i) letter changes of s are *alphabetic variants* of s; and (ii) if β' is an alphabetic variant of β, then $\sigma x.\beta'$ is an *alphabetic variant* of φ.

It is easy to see that if M and N are well-formed expressions and M is an alphabetic variant of N then either M and N are both σ-terms or both are formulas, and that $\mathbf{FV}(M) = \mathbf{FV}(N)$. Indeed, the only difference between the two is that some bound occurrences of variables from M may have been replaced by bound occurrences of other variables in N. It is also readily checked that the relation of being alphabetic variants is an equivalence relation.

Obviously since we have an infinite stock of variables we can always find, for any finite set $\{y_1, \ldots, y_n\} \subseteq \mathbf{Var}$ and each formula φ or term s, an alphabetic variant φ' or s' such that none of $\{y_1, \ldots, y_n\}$ occurs bound in φ' or s'. We now set out a systematic method of doing this in a way which guarantees, for a given term t and variable x, that t is free for x in φ' or s'.

Definition 6. For each well-formed expression M we define M' recursively by

(1) φ is atomic and σ-free $\implies$ $\varphi \equiv \varphi'$.

(2) $\varphi \equiv \neg\beta$ $\implies$ $\varphi' \equiv \neg\beta'$.

(3) $\varphi \equiv \psi \vee \beta$, $\varphi \equiv \psi \wedge \beta$, or $\varphi \equiv \psi \Rightarrow \beta$ $\implies$ $\varphi' \equiv \psi' \vee \beta'$, $\varphi' \equiv \psi' \wedge \beta'$, or $\varphi' \equiv \psi' \Rightarrow \beta'$.

(4) $\varphi \equiv \forall y.\beta$ or $\varphi \equiv \exists y.\beta$ $\implies$ (i) if $x \notin \mathbf{FV}(\varphi)$, then $\varphi' \equiv \varphi$; (ii) $x \in \mathbf{FV}(\varphi)$ and $y \notin \mathbf{FV}(t)$, then $\varphi' \equiv \forall y.\beta'$; (iii) $x \in \mathbf{FV}(\varphi)$ and $y \in \mathbf{FV}(t)$, then $\varphi' \equiv \forall z.(\beta[[y/z]])$ or $\varphi' \equiv \exists z.(\beta[[y/z]])$, where z is the first variable which does not occur free in t and such that φ' is a letter change for $\forall y.\beta'$ or $\exists y.\beta'$ (i. e. z is not free in β', but is free for y in β').

(5) $s \equiv \sigma y.\beta$ $\implies$ (i) if $x \notin \mathbf{FV}(s)$, then $s' \equiv s$; (ii) $x \in \mathbf{FV}(s)$ and $y \notin \mathbf{FV}(t)$, then $s' \equiv \sigma y.\beta'$; (iii) $x \in \mathbf{FV}(s)$ and $y \in \mathbf{FV}(t)$, then $s' \equiv \sigma z.(\beta[[y/z]])$, where z is the first variable which does not occur free in t and such that s' is a letter change for $\sigma y.\beta'$.

(6) φ is atomic and not σ-free $\implies$ (i) if $\varphi \equiv Pt_1 \dots t_n$ ($\varphi \equiv t_1 = t_2$) and $s_1, \dots, s_m$ are all the occurrences of σ-terms among $t_1, \dots, t_n$ (t_1, t_2), then $\varphi' \equiv \varphi[[s_1/s'_1]] \dots [[s_n/s'_n]]$.

Note that in the last clause we can get away with consecutive substitution (i. e. we don't need simultaneous substitution) because we are not replacing embedded σ–terms. Note that s' is an alphabetic variant of s and φ' is an alphabetic variant of φ. Note also that if t is free for x in φ (in s), then $\varphi' \equiv \varphi$ ($s' \equiv s$).

Finally we are ready to give our definition of substitution.

Definition 7 (Substitution). Let $s, s_1, \dots, s_n \in$ **Terms**, $\varphi, \beta, \psi \in$ **Wffs**, $x \in$ **Var**, and $t \in$ **Terms**. Then $s[x/t]$ and $\varphi[x/t]$, the *result of (properly) substituting an occurrence of t for each free occurrence of x in s (in φ)* are defined recursively as follows.

(1) $s \equiv x \implies s[x/t] \equiv t$.

(2) $s \equiv y \not\equiv x \implies s[x/t] \equiv y$.

(3) $\varphi \equiv Ps_1 \dots s_n \implies \varphi[x/t] \equiv Ps'_1[x/t] \dots s'_n[x/t]$.

(4) $\varphi \equiv s_1 = s_2 \implies \varphi[x/t] \equiv s'_1[x/t] = s'_2[x/t]$.

(5) $\varphi \equiv \neg\beta \implies \varphi[x/t] \equiv \neg\beta'[x/t]$.

(6) $\varphi \equiv \psi \vee \beta$, $\varphi \equiv \psi \wedge \beta$, or $\varphi \equiv \psi \Rightarrow \beta \implies \varphi[x/t] \equiv \psi'[x/t] \vee \beta'[x/t]$, $\varphi[x/t] \equiv \psi'[x/t] \wedge \beta'[x/t]$ or $\varphi[x/t] \equiv \psi'[x/t] \Rightarrow \beta'[x/t]$

(7) $\varphi \equiv \forall y.\beta$ or $\varphi \equiv \exists y.\beta \implies$ (i) if $x \notin$ **FV**(φ), then $\varphi[x/t] \equiv \varphi'(\equiv \varphi)$; (ii) if $x \in$ **FV**(φ) and t is free for x in φ, then $\varphi[x/t] \equiv \forall y.(\beta[x/t])$ or $\varphi[x/t] \equiv \exists y.(\beta[x/t])$; (iii) if $x \in$ **FV**(φ) and t is not free for x in φ, then

$\varphi[x/t] \equiv \varphi'[x/t]$.

(8) $s \equiv \sigma y.\beta \implies$ (i) if $x \notin \mathbf{FV}(s)$, then $s[x/t] \equiv s'(\equiv s)$; (ii) if $x \in \mathbf{FV}(s)$ and t is free for x in s, then $s[x/t] \equiv \sigma y.(\beta[x/t])$; (iii) if $x \in \mathbf{FV}(s)$ and t is not free for x in s, then $s[x/t] \equiv s'[x/t]$.

Note that in the last two clauses we could have directly paralleled the corresponding clause of the definition of substitution in the λ-calculus by writing, for instance, for the clauses (ii) and (iii) of (7), that $\varphi[x/t] \equiv \forall z.(\beta[y/z][x/t])$ where $z \equiv y$ if t is free for x in φ, otherwise z is the first variable not free in β but free for x in β.

In our treatment of the λ-calculus, after defining substitution we adopted Variable Convention I—that is, we declared that we would work with α-equivalence classes of λ-terms rather than the λ-terms themselves. This convention made sense in part because in a language for the λ-calculus all the well-formed expressions are of the same sort, namely λ-terms. Since we now must take into account the formation of terms out of formulas and the substitution of terms into formulas, a similar move would require also treating formulas as representatives of some analogue of α-equivalence classes. While this would be possible, it would result in a rather non-standard treatment of first order logic. In logic it is usual to prove that alphabetic variants are logically equivalent (or provably equivalent, depending on the approach preferred by the author). We will simply extend this result to take terms into account as well as formulas. (More precisely, we will adopt rules sufficient to allow us to show that terms which are alphabetic variants are provably equal.)

However, there is still one step we need to take before quitting with substitution. It will be useful to have a notion of simultaneous substitution.

Definition 8. Let M be a term or a formula, $\vec{x} = \langle x_1, \ldots, x_n \rangle$ be a sequence of distinct variables, and $\vec{t} = \langle t_1, \ldots, t_n \rangle$ a sequence of terms. The *simultaneous substitution of $\vec{t}$ for $\vec{x}$ in M*, which is written either as '$M[x_1/t_1, \ldots, x_n/t_n]$' or as '$M[\vec{x}/\vec{t}]$', is defined by recursion on n. The definition of substitution gives us the $n = 1$ case. For $n > 1$, if $x_n \notin \mathbf{FV}(M)$,

$$M[\vec{x}/\vec{t}] \equiv M[x_1/t_1, \ldots, x_{n-1}/t_{n-1}].$$

For $x_n \in \mathbf{FV}(M)$, let z be the first variable in neither M nor in any of $t_1, \ldots, t_n$. Note that by the inductive hypothesis we know how to get

$$M[x_1/t_1[x_n/z], \ldots, x_{n-1}/t_{n-1}[x_n/z]].$$

So we define

$$M[\vec{x}/\vec{t}] \equiv M[x_1/t_1[x_n/z], \ldots, x_{n-1}/t_{n-1}[x_n/z]][x_n/t_n][z/x_n].$$

In effect, we use z as a placeholder which prevents t_n from being plugged in at inappropriate places in the other terms placed in M already, so we can get the effect of simultaneous substitution by repeated application of ordinary substitution.

3. The Systems

We now need to introduce the rules and axioms for the various logical systems we will be dealing with. We begin with the

Propositional Axioms[2]**.**

(P1) $\alpha \Rightarrow (\beta \Rightarrow \alpha)$

(P2) $(\alpha \Rightarrow (\beta \Rightarrow \gamma)) \Rightarrow ((\alpha \Rightarrow \beta) \Rightarrow (\alpha \Rightarrow \gamma))$

(P3) $\alpha \Rightarrow (\beta \Rightarrow (\alpha \wedge \beta))$

(P4) $(\alpha \wedge \beta) \Rightarrow \alpha$

(P5) $(\alpha \wedge \beta) \Rightarrow \beta$

(P6) $\alpha \Rightarrow (\alpha \vee \beta)$

(P7) $\beta \Rightarrow (\alpha \vee \beta)$

(P8) $(\alpha \Rightarrow \gamma) \Rightarrow ((\beta \Rightarrow \gamma) \Rightarrow ((\alpha \vee \beta) \Rightarrow \gamma))$

(P9) $(\alpha \Rightarrow \beta) \Rightarrow ((\alpha \Rightarrow \neg\beta) \Rightarrow \neg\alpha)$

(P10) $\neg\alpha \Rightarrow (\alpha \Rightarrow \beta)$

(P11) $\alpha \vee \neg\alpha$

Definition 9. The set of *axioms of the intuitionistic propositional calculus* is the smallest set such that, for any $\alpha, \beta, \gamma \in$ **Wffs**, (P1)-(P10) are in the set. The set of *axioms of the classical propositional calculus* is the smallest set such that, for any $\alpha, \beta, \gamma \in$ **Wffs**, (P1)-(P11) are in the set.

Quantification Axioms[3]**.** For any $\alpha, \beta \in$ **Wffs**, the following are quantificational axioms.

(Q1) $\forall x.(\alpha \Rightarrow \beta) \Rightarrow (\forall x.\alpha \Rightarrow \forall x.\beta)$

(Q2) $\forall x.(\alpha \Rightarrow \beta) \Rightarrow (\exists x.\alpha \Rightarrow \exists x.\beta)$

[2] These axioms are the propositional axioms for intuitionistic logic from [Bell and Machover] plus excluded middle, and they are, of course, equivalent to any of the usual axiomatizations. This is certainly not the most efficient way to axiomatize the classical case, but we will use this set of axioms in the name of uniformity of treatment.

[3] These, too, contain a good deal of redundancy in the classical case.

(Q3) $\alpha \Rightarrow \forall x.\alpha$ provided x is not free in α.

(Q4) $\exists x.\alpha \Rightarrow \alpha$ provided x is not free in α.

(Q5) $\forall x.\alpha \Rightarrow \alpha[x/t]$ provided t is free for x in α.

(Q6) $\alpha[x/t] \Rightarrow \exists x.\alpha$ provided t is free for x in α.

Definition 10. If $\alpha \in$ **Wffs**, a *generalization of* α is a formula of the form $\forall x_1 \ldots \forall x_k.\alpha$, where $k \geqslant 0$ and $x_1 \ldots x_k \in$ **Var** are not necessarily distinct. So, in particular, α is a generalization of itself. The set of *axioms of the intuitionistic first order predicate calculus* is the smallest set including all generalizations of the intuitionistic propositional axioms and all generalizations of quantificational axioms. The set of *axioms of the classical first order predicate calculus* is the smallest set including all generalizations of the classical propositional axioms and all generalizations of quantificational axioms.

Axioms of Identity. Let $t_1, \ldots, t_{2n} \in$ **Terms**, and $P \in$ **Pred**$_n$. The following are axioms of identity.

(I1) $t_1 = t_1$.

(I2) $(t_1 = t_2 \wedge t_2 = t_3) \Rightarrow t_1 = t_3$.

(I3) $(t_1 = t_{n+1} \wedge \cdots \wedge t_n = t_{2n} \wedge Pt_1 \ldots t_n) \Rightarrow Pt_{n+1} \ldots t_{2n}$.

Definition 11. The set of *axioms of the intuitionistic first order predicate calculus with identity* is the union of the set of axioms of the intuitionistic first order predicate calculus and the set of generalizations of identity axioms. The set of *axioms of the classical first order predicate calculus with identity* is the union of the set of axioms of the classical first order predicate calculus and the set of generalizations of identity axioms.

Rules of Inference. For the most part we will need only *modus ponens*.

(MP)
$$\frac{\alpha \qquad \alpha \Rightarrow \beta}{\beta}$$

However, we sometimes also will look, for each operator σ,[4] at

($\sigma*$)
$$\frac{\forall x.(\varphi \Leftrightarrow \psi)}{\sigma x.\varphi = \sigma x.\psi}.$$

Definition 12. Each of the *classical propositional calculus*, the *intuitionistic propositional calculus*, the *classical first order predicate calculus*, the *intuitionistic first order predicate calculus*, the *classical first order predicate calculus with identity*, and the *intuitionistic first order predicate calculus with identity*, is defined to be the calculus which results by taking as axioms the corresponding set as defined above, and *modus ponens* as the sole rule of inference.

The following are axioms which have versions for each operator.

Axioms for the σ–terms. Let $\varphi, \psi \in$ **Wffs**, $t_1, t_2 \in$ **Terms**.

(σext) $\forall x.(\varphi \Leftrightarrow \psi) \Rightarrow \sigma x.\varphi = \sigma x.\psi$.

(σA6) $t_1 = t_2 \Rightarrow \sigma x.\varphi[y/t_1] = \sigma x.\varphi[y/t_2]$.

($\sigma\alpha$) $\sigma x.\varphi = \sigma y.\varphi[x/y]$

The last of these is obviously named for the α–rule we encountered in our discussion of the λ-calculus. We will include this axiom schema in all the systems below. This will allow us to show that terms that are alphabetic variants are provably equal.

[4] We will also have occasion to look at some modified versions of this rule in Chapters 9, 11 and 12.

The (σext) and (σA6) axioms and the (*) rule of inference are all obviously various types of extensionality requirements that we might put on the behaviour of σ–terms. In general, we will call a calculus for a particular operator *extensional* if it satisfies the (σext) axioms; we will call it the σ^* calculus if the (*) rule is valid; and we will call it (for historical reasons to be mentioned below) the Hilbert–Bernays σ–calculus if it satisfies (A6).

Not surprisingly, when we do not include axioms which guarantee that terms will behave extensionally some of the rules normally used in first order logic no longer hold in full generality (for instance, the substitutivity of identicals will sometimes fail). We will therefore introduce the following modified versions of two of our axioms.

(Q5′) $\forall x.\alpha \Rightarrow \alpha[x/t]$ provided t is free for x in α, and provided x does not have a free occurrence in the scope of an σ–bound variable.

(Q6′) $\alpha[x/t] \Rightarrow \exists x.\alpha$ provided t is free for x in α, and provided x does not have a free occurrence in the scope of an σ–bound variable.

The various operators which we will give their own name will be distinguished from one another by their characteristic axioms, which is to say that they are distinguished from one another by their logical behaviour. We will introduce here only the two operators which will occupy us in Chapters 7–11. We will introduce some other in Chapter 12.

Characteristic Axioms for ε and τ. Let $\varphi \in$ **Wffs**. Then the following are axioms.

(ε) $\exists x.\varphi \Rightarrow \varphi[x/\varepsilon x.\varphi]$.

(τ) $\varphi[x/\tau x.\varphi] \Rightarrow \forall x.\varphi$.

Definition 13. As an unstated assumption, all the various sets of axioms described in this paragraph include all instances of the ($\varepsilon\alpha$)-schema. The set of *axioms of the intuitionistic ε-calculus* is the union of the set of axioms of the intuitionistic first order predicate calculus with identity, where we use (Q5′) and (Q6′) instead of (Q5) and (Q6), and the set of generalizations of (ε)-axioms. Similarly for the *axioms of the classical ε-calculus*. If we use the original versions of (Q5) and (Q6) and we also add the generalizations of axioms of the form (εext), we will call the result the set of axioms of the *extensional* (intuitionistic or classical) ε-calculus. If rather than the (εext) axioms we add the generalizations of the axioms of form (A6), the result is called the set of axioms of the (intuitionistic or classical) *Hilbert-Bernays* ε-calculus.[5] The *classical ε-calculus*, the *intuitionistic ε-calculus*, the *classical extensional ε-calculus*, the *intuitionistic extensional ε-calculus*, the *classical Hilbert-Bernays ε-calculus*, and the *intuitionistic Hilbert-Bernays ε-calculus* are each defined to be the calculus which results by taking as axioms the corresponding set as defined as above and *modus ponens* as the sole rule of inference. On the other hand, the *classical* and *intuitionistic ε^*-calculi* are formed by taking as axioms the set of axioms of the classical or intuitionistic ε-calculus, and taking as rules of inference *modus ponens* and ($\varepsilon*$).

---

[5] Another approach we could have taken to this is to adopt (Q5′) and (Q6′) for all the systems. A straightforward inductive argument shows that if we have (σA6) we can prove the general versions (Q5) and (Q6), and (σext) obviously implies (σA6), so the result would be the same as the one we get by taking the approach described in this chapter.

We can define all the various sets of axioms and calculi for τ by rewriting the last paragraph with τ in place of ε.

Comment. We will call (εext) '(Ack)' after Wilhelm Ackermann, who is sometimes credited with being the first to formulate it. Axiom (εA6) is from the formalization of the ε-calculus in [Hilbert and Bernays, vol. 2], (which is why we call calculi which incorporate rules of this form what we do) and seems to be a formal version of some comments made about reasoning with ε-terms at the end of [Hilbert 1927]. The (*) rules are investigated here for the first time, and are interesting primarily in the ε case, i. e. as a contrast to (Ack).

CHAPTER VI

CLASSICAL FIRST ORDER SEMANTICS

The material in this brief chapter will no doubt be very familiar to its readers. The point of including it is two-fold. First, it will make explicit the terminology and the formulations of key definitions for first order semantics that we will be using here. Secondly, by stating (without proof) some of the basic results of first order model theory we will be more easily able to refer to those results when we need them later.

If $\mathcal{L}$ is a language as defined in Chapter 5, we can recover a standard language for the first order predicate calculus with equality by taking the subset of the well-formed expressions $\mathcal{L}' = \{ M \mid M \text{ is } \sigma\text{-free} \}$. We assume that the predicate symbols and constants of $\mathcal{L}'$ are given in the form of indexed sets $\{ P_i \mid i \in I \}$ and $\{ c_j \mid j \in J \}$, and for each $i \in I$ we denote the arity of P_i by $\zeta(i)$. The map ζ from I to the set of positive integers is called the *signature* of $\mathcal{L}'$.

An $\mathcal{L}'$ structure is an ordered triple $\mathcal{M} = \langle D, \mathcal{R}, \mathbf{c} \rangle$ where

(1) $D \neq \emptyset$ is a set

(2) $\mathcal{R} : I \to \{ f \mid \text{ for some positive integer } n,\ f : D^n \to \{0,1\}\}$ is such that for each $i \in I$, $\mathcal{R}_i : D^{\zeta(i)} \to \{0,1\}$

(3) $\mathbf{c} : J \to D$.

We define a valuation in $\mathcal{M}$ to be a map $\varrho : \mathbf{Var} \to D$. If ϱ is a valuation, the valuation which agrees with ϱ except that it assigns $d \in D$ to v_n is defined to be the map $\varrho(v_n/d) : \mathbf{Var} \to D$ such that

$$\varrho(v_n/d) = \begin{cases} \varrho(v_i) & \text{for } i \neq n \\ d & \text{if } i = n. \end{cases}$$

For later convenience we will regard the set $\{0, 1\}$ as the Boolean algebra $\mathbf{2}$.

Definition 1. We define for each term t and each formula φ of $\mathcal{L}'$ the *interpretations, $[t]^{\varrho}_{\mathcal{M}} \in D$ and $\|\varphi\|^{\varrho}_{\mathcal{M}} \in \mathbf{2}$, of t and φ in $\mathcal{M}$ under ϱ* recursively as follows.

$$[x]^{\varrho}_{\mathcal{M}} = \varrho(x) \qquad \text{for } x \in \mathbf{Var}$$

$$[c_i]^{\varrho}_{\mathcal{M}} = \mathbf{c}(i) \qquad \text{for } c_i \in \mathbf{Con}$$

$$\|P_i t_1, \ldots, t_{\zeta(i)}\|^{\varrho}_{\mathcal{M}} = \mathcal{R}_i\left(\langle [t_1]^{\varrho}_{\mathcal{M}}, \ldots, [t_{\zeta(i)}]^{\varrho}_{\mathcal{M}} \rangle\right)$$

$$\|t_1 = t_2\|^{\varrho}_{\mathcal{M}} = \begin{cases} 1 & \text{if } [t_1]^{\varrho}_{\mathcal{M}} = [t_2]^{\varrho}_{\mathcal{M}} \\ 0 & \text{otherwise} \end{cases}$$

$$\|\neg\varphi\|^{\varrho}_{\mathcal{M}} = (\|\varphi\|^{\varrho}_{\mathcal{M}})^*$$

$$\|\varphi \wedge \psi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho}_{\mathcal{M}} \wedge \|\psi\|^{\varrho}_{\mathcal{M}}$$

$$\|\varphi \vee \psi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho}_{\mathcal{M}} \vee \|\psi\|^{\varrho}_{\mathcal{M}}$$

$$\|\varphi \Rightarrow \psi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho}_{\mathcal{M}} \Rightarrow \|\psi\|^{\varrho}_{\mathcal{M}} \quad (= (\|\varphi\|^{\varrho}_{\mathcal{M}})^* \vee \|\psi\|^{\varrho}_{\mathcal{M}})$$

$$\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} = \bigvee_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}}$$

$$\|\forall x.\varphi\|^{\varrho}_{\mathcal{M}} = \bigwedge_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}}.$$

The symbols $\Rightarrow$, $\vee$, $\wedge$, and * on the right denote the operations in the Boolean algebra $\mathbf{2}$. It is easily checked that these definitions are precisely the same as the usual ones given in English, where we take 1 to designate truth and 0 falsity. For instance, the clause for the existential quantifier tells us that $\exists x.\varphi$ is satisfied by ϱ if and only if $\varrho(x/d)$ satisfies φ for some $d \in D$. If $\|\varphi\|^{\varrho}_{\mathcal{M}} = 1$ for some $\varrho : \mathbf{Var} \to D$, we say that φ is *satisfiable* (in $\mathcal{M}$), and that it is *satisfied by* ϱ. If $\|\varphi\|^{\varrho}_{\mathcal{M}} = 1$ for all $\varrho : \mathbf{Var} \to D$, we say that $\mathcal{M}$ is a *model* for φ, and write $\mathcal{M} \models_{PC} \varphi$. If $\mathcal{M} \models_{PC} \varphi$ for all $\mathcal{L}'$ structures $\mathcal{M}$, we say that φ is *valid in the classical first order predicate calculus*, and we write '$\models_{PC} \varphi$'. If Σ is a set of formulas, we say that Σ is *satisfiable* if there is some ϱ in some $\mathcal{M}$ that satisfies every member of Σ. We say that $\mathcal{M}$ is a model of Σ if $\mathcal{M}$ is a model of every formula in Σ, and we write '$\mathcal{M} \models_{PC} \Sigma$' for this. We write '$\Sigma \models_{PC} \varphi$' if

$$\mathcal{M} \models_{PC} \Sigma \implies \mathcal{M} \models_{PC} \varphi.$$

We will write *card*($\mathcal{M}$) for the cardinality of D. We will also write *card*($\mathcal{L}'$) for the cardinality of the set of well-formed expressions of $\mathcal{L}'$. This is always $\aleph_0$ as we have defined languages. If we remove the restrictions on the maximum cardinality of **Pred** and **Con**, it can of course be larger. It is this that gives point to, for example, the strong completeness theorem stated below.

We will state without proof several well-known results about the classical predicate calculus. Proofs can be found in most logic textbooks, for instance in [Bell and Machover, ch. 3 and ch. 5]. We write '$\Sigma \vdash_{PC} \varphi$' if φ is derivable from Σ in the first order predicate calculus with identity.

Theorem 1 (Soundness Theorem for the Classical Predicate Calcu-

lus).

$$\Sigma \vdash_{PC} \varphi \implies \Sigma \models_{PC} \varphi$$

□

We state a variety of completeness theorems. First we state the version of the completeness theorem we can get without appeal to either the Axiom of Choice or the Ultrafilter Theorem.

Theorem 2 (Gödel's Completeness Theorem).

(1) *Every countable consistent set of formulas of $\mathcal{L}'$ is satisfiable in some structure with a countable universe.*

(2) *If Σ is countable, $\Sigma \models_{PC} \varphi \implies \Sigma \vdash_{PC} \varphi$.* □

We have defined $\mathcal{L}'$ in such a way that the restriction to countable sets of formulas is in fact redundant, but by removing this restriction we give point to

Theorem 3 (Gödel–Henkin Completeness Theorem). *Σ is satisfiable if and only if Σ is consistent.* □

We state also the familiar

Theorem 4 (Compactness Theorem). *If every finite subset of Σ is satisfiable, then Σ is satisfiable.* □

The next theorem is proved in [Bell and Slomson, pp. 103–106].

Theorem 5. *Theorem 3, Theorem 4 and Theorem 2.1 are all equivalent.* □

If we are willing to appeal to the Axiom of Choice we can prove strong versions of the Löwenhein–Skolem Theorem, for instance

Theorem 6 (Löwenhein–Skolem Theorem). *If* Σ *is satisfiable in* $\mathcal{M}$ *and* $card(\mathcal{M}) \geqslant \aleph_0$ *then, for every cardinal* $\kappa \geqslant max(card(\Sigma), \aleph_0)$, Σ *is satisfiable in some* $\mathcal{M}'$ *such that* $card(\mathcal{M}') = \kappa$. □

Indeed, it is also shown by Bell and Slomson that the Axiom of Choice is equivalent to the

Theorem 7 (Strong Completeness Theorem). *Let* κ *be an infinite cardinal. If* Σ *is consistent and* $card(\Sigma) \leqslant \kappa$, *then either* Σ *is satisfiable in a finite structure or* Σ *is satisfiable in some* $\mathcal{M}'$ *with* $card(\mathcal{M}') = \kappa$. □

CHAPTER VII

SEMANTICS FOR THE EXTENSIONAL CLASSICAL ε—CALCULUS

When in the past people have turned their attention to semantics for the ε operator—or, indeed, as we shall see in Chapter 12, when they have turned their attention to semantics for term-forming operators in general—they have usually dealt with systems which include the extensionality principle we called (σext) in Chapter 5. The reason for this is not hard to see. For if we are to give the most straightforward possible interpretation to the ε-terms, the extensionality principle is required for completeness to be provable.

In this chapter we present a semantics for the extensional classical ε-calculus. The treatment is essentially the same as the one found in [Leisenring] and to the first semantics presented in [Asser], except that we will need to generalize things so that we can handle open terms. We will present detailed proofs of soundness and completeness, and of a few other famous results of the (classical) ε-calculus. The point of doing so is partly that no earlier proof in the literature handles open terms, and partly that giving these proofs in detail will allow us to describe some later proofs in terms of the modifications which need to be made to the ones in this Chapter.

1. The Machinery

We can easily build a semantics for the classical extensional ε-calculus by interpreting the ε-terms using choice functions on the power sets of the domains of the classical structures described in Chapter 6.

Let $\mathcal{L}$ be a language for the ε-calculus. $\mathcal{M} = \langle D, \mathcal{R}, \mathbf{c}, E\rangle$ is a (classical, extensional) ε-structure for $\mathcal{L}$ if $\langle D, \mathcal{R}, \mathbf{c}\rangle$ is a structure for $\mathcal{L}' \subseteq \mathcal{L}$, the ε-free portion of $\mathcal{L}$ as described in Chapter 6, and $E = E' \cup \{\langle \emptyset, m\rangle\}$ where E' is a choice function on $\mathcal{P}(D)$ and m is an arbitrary but fixed element of D. (So for $X \subseteq D$, $X \neq \emptyset \implies E(X) \in X$, and $E(\emptyset) = m$.)

The definitions for $\varrho, \varrho(x/d) : \mathbf{Var} \to D$ are as before.

Definition 1. To extend the definitions of $[t]^{\varrho}_{\mathcal{M}} \in D$ and $\|\varphi\|^{\varrho}_{\mathcal{M}} \in \mathbf{2}$ from the terms and formulas of $\mathcal{L}'$ to all of **Terms** and **Wffs**, we can rewrite definition 6.1, except that we write $\mathcal{L}$ for $\mathcal{L}'$, and we add the following clause to it:

$$[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = E\left(\left\{ d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 1\right\}\right).$$

We pause to show that our machinery is working properly.

Theorem 1. *Let $\mathcal{M}$ be a classical extensional ε-structure for $\mathcal{L}$, M a well-formed expression of $\mathcal{L}$, and ϱ, ϱ' two valuations in $\mathcal{M}$ such that for any $x \in \mathbf{FV}(M)$, $\varrho(x) = \varrho'(x)$. Then*

(1) $M \equiv t \in \mathbf{Terms} \implies [t]^{\varrho}_{\mathcal{M}} = [t]^{\varrho'}_{\mathcal{M}}$

(2) $M \equiv \varphi \in \mathbf{Wffs} \implies \|\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho'}_{\mathcal{M}}$.

Proof. By induction on comp(M). The result is trivial for variables and constants. So suppose $[t_j]^{\varrho}_{\mathcal{M}} = [t_j]^{\varrho'}_{\mathcal{M}}$ for $j = 1, \ldots, n$ and $P_i \in \mathbf{Pred}_n$. Then

$$\begin{aligned} \|P_i t_1 \ldots t_n\|^{\varrho}_{\mathcal{M}} &= \mathcal{R}_i\left(\langle [t_1]^{\varrho}_{\mathcal{M}}, \ldots, [t_n]^{\varrho}_{\mathcal{M}} \rangle\right) \\ &= \mathcal{R}_i\left(\left\langle [t_1]^{\varrho'}_{\mathcal{M}}, \ldots, [t_n]^{\varrho'}_{\mathcal{M}} \right\rangle\right) \\ &= \|P_i t_1 \ldots t_n\|^{\varrho'}_{\mathcal{M}}. \end{aligned}$$

The clause for $=$ is similar, and the clauses for $\neg$, $\wedge$, $\vee$ and $\Rightarrow$ are routine.

So suppose $\|\varphi\|^{\varrho_1}_{\mathcal{M}} = \|\varphi\|^{\varrho_2}_{\mathcal{M}}$ for any ϱ_1, ϱ_2 which agree on $\mathbf{FV}(\varphi)$. Since ϱ and ϱ' agree on $\mathbf{FV}(\exists x.\varphi)$ ($= \mathbf{FV}(\forall x.\varphi) = \mathbf{FV}(\varepsilon x.\varphi)$), for any $d \in D$, $\varrho(x/d)$ and $\varrho'(x/d)$ agree on $\mathbf{FV}(\varphi)$. So by inductive hypothesis,

$$\|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \|\varphi\|^{\varrho'(x/d)}_{\mathcal{M}}$$

for all $d \in D$. Thus

$$\bigvee_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d \in D} \|\varphi\|^{\varrho'(x/d)}_{\mathcal{M}},$$

i. e. $\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} = \|\exists x.\varphi\|^{\varrho'}_{\mathcal{M}}$. Similarly for $\forall$.

It also follows that $\{ d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 1 \} = \{ d \in D \mid \|\varphi\|^{\varrho'(x/d)}_{\mathcal{M}} = 1 \}$, so E will choose the same element from either set, and so $[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = [\varepsilon x.\varphi]^{\varrho'}_{\mathcal{M}}$. □

Of course, it follows as usual that the value assigned to any term or formula depends only on the value assigned to its free variables, and so the interpretation of a closed term or sentence in a structure is independent of ϱ. So we will sometimes write '$\|M\|_{\mathcal{M}}$' (or '$[M]_{\mathcal{M}}$') for $\|M\|^{\varrho}_{\mathcal{M}}$ ($[M]^{\varrho}_{\mathcal{M}}$) if M is closed.

It remains to check that the machinery of substitution works appropriately. It is here that we must pay the price, namely a rather tedious bookkeeping proof, for having enough moral fibre not to simply restrict substitution to cases where the substituting term is free for the substituted variable and restrict our rules accordingly. This would not have eliminated the need for Theorem 2, but it might have made its proof less tedious.

Theorem 2. *Let M be a well-formed expression of $\mathcal{L}$, $\mathcal{M}$ a structure for $\mathcal{L}$, ϱ a valuation in $\mathcal{M}$, and $s \in$* **Terms**. *Then*

(1) $M \equiv t \in \textbf{Terms} \implies [t]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = [t[x/s]]_{\mathcal{M}}^{\varrho}$

(2) $M \equiv \varphi \in \textbf{Wffs} \implies \|\varphi\|_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = \|\varphi[x/s]\|_{\mathcal{M}}^{\varrho}$.

Proof. If $x \notin \mathbf{FV}(M)$, $M[x/s] \equiv M$, so the result follows from the last theorem. So we assume $x \in \mathbf{FV}(M)$. First we prove

Lemma 1. *If s is free for x in M, the theorem holds.*

Proof of Lemma 1. The proof is by induction on comp(M), following the list of cases in the definition of "s is free for x" (Definition 5.4), ignoring the cases where x is not free in M.

First, $[x]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = [s]_{\mathcal{M}}^{\varrho}$ by the definition of $[\cdot]$.

So suppose that for $i = 1, \ldots, n$, $[t_i]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = [t_i[x/s]]_{\mathcal{M}}^{\varrho}$. Then

$$\begin{aligned}
\|P_i t_1 \ldots t_n\|_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} &= \mathcal{R}_i \left(\left\langle [t_1]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})}, \ldots, [t_n]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} \right\rangle\right) \\
&= \mathcal{R}_i(\langle [t_1[x/s]]_{\mathcal{M}}^{\varrho}, \ldots, [t_n[x/s]]_{\mathcal{M}}^{\varrho} \rangle) \\
&= \|(P_i t_1 \ldots t_n)[x/s]\|_{\mathcal{M}}^{\varrho}
\end{aligned}$$

with the last equality justified since s is free for x in each t_i, so $t_i' \equiv t_i$, and by the definition of substitution $(P_i t_1 \dots t_n)[x/s] \equiv P_i t'_1[x/s] \dots t'_n[x/s]$. The $=$ case is similar.

Since $\varphi \equiv \varphi'$ when s is free for x in φ, the cases of $\neg$, $\wedge$, $\vee$, and $\Rightarrow$ are entirely routine.

As we are ignoring the case where $x \notin \mathbf{FV}(M)$, suppose s is free for x in β, y does not occur in s, and $x \not\equiv y$. By the induction hypothesis, for any valuation, and so for each valuation $\varrho(y/d)$ with $d \in D$,

$$\|\beta\|_{\mathcal{M}}^{\varrho(y/d)(x/[s]^{\varrho}_{\mathcal{M}})} = \|\beta[x/s]\|_{\mathcal{M}}^{\varrho(y/d)}.$$

Since $x \not\equiv y$, this yields

$$\|\beta\|_{\mathcal{M}}^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})(y/d)} = \|\beta[x/s]\|_{\mathcal{M}}^{\varrho(y/d)}$$

for each $d \in D$, and so

$$\{\, d \in D \mid \|\beta\|_{\mathcal{M}}^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})(y/d)} = 1 \,\} = \{\, d \in D \mid \|\beta[x/s]\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \,\}.$$

Now it follows, using Definition 1, that $\|\forall y.\beta\|_{\mathcal{M}}^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})} = \|\forall y.\beta[x/s]\|_{\mathcal{M}}^{\varrho}$, that $\|\exists y.\beta\|_{\mathcal{M}}^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})} = \|\exists y.\beta[x/s]\|_{\mathcal{M}}^{\varrho}$, and that $[\varepsilon y.\beta]_{\mathcal{M}}^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})} = [\varepsilon y.\beta[x/s]]_{\mathcal{M}}^{\varrho}$. This completes the proof of Lemma 1. □

We also will need

Lemma 2. *If M' arises out of M by a letter change, then for any ϱ, $\mathcal{M}$*

(1) $M \in \mathbf{Terms} \implies [M']^{\varrho}_{\mathcal{M}} = [M]^{\varrho}_{\mathcal{M}}$

(2) $M \in \mathbf{Wffs} \implies \|M'\|^{\varrho}_{\mathcal{M}} = \|M\|^{\varrho}_{\mathcal{M}}$.

Proof of lemma 2. M has one of the forms $\forall x.\beta$, $\exists x.\beta$, or $\varepsilon x.\beta$, while z is not free in β but is free for x in β, thus M' is $\forall z.\beta[x/z]$, or $\exists z.\beta[x/z]$ or $\varepsilon z.\beta[x/z]$. By lemma 1, for any $d \in D$

$$\|\beta[x/z]\|_{\mathcal{M}}^{\varrho(z/d)} = \|\beta\|_{\mathcal{M}}^{\varrho(z/d)(x/[z]_{\mathcal{M}}^{\varrho(z/d)})} = \|\beta\|_{\mathcal{M}}^{\varrho(z/d)(x/d)} = \|\beta\|_{\mathcal{M}}^{\varrho(x/d)}$$

using the fact that z is not free in β. So

$$\{\, d \in D \mid \|\beta[x/z]\|_{\mathcal{M}}^{\varrho(z/d)} = 1 \,\} = \{\, d \in D \mid \|\beta\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \,\}.$$

The results follow. $\square$

The next lemma that we need takes longer to state than to prove.

Lemma 3. *Suppose that for all ϱ, $\mathcal{M}$, $[t]_{\mathcal{M}}^{\varrho} = [t']_{\mathcal{M}}^{\varrho}$ and $\|\beta\|_{\mathcal{M}}^{\varrho} = \|\beta'\|_{\mathcal{M}}^{\varrho}$.*

(1) *If t is a subterm of s, and s' results from s by replacing 0 or more occurrences of t by t', then for any ϱ, $\mathcal{M}$, $[s]_{\mathcal{M}}^{\varrho} = [s']_{\mathcal{M}}^{\varrho}$.*

(2) *If t occurs in φ, and φ' results from φ by replacing 0 or more occurrences of t by t', then for any ϱ, $\mathcal{M}$, $\|\varphi\|_{\mathcal{M}}^{\varrho} = \|\varphi'\|_{\mathcal{M}}^{\varrho}$.*

(3) *If β is a subformula of φ, and φ' results from φ by replacing 0 or more occurrences of β by β', then for any ϱ, $\mathcal{M}$, $\|\varphi\|_{\mathcal{M}}^{\varrho} = \|\varphi'\|_{\mathcal{M}}^{\varrho}$.*

Proof of lemma 3. An inductive argument on the complexity of φ and t which proceeds up the steps in Definition 5.2 does the trick for all three parts. The proof is an entirely routine application of Definition 1. We do the final case because it will fail in an interesting way when we give non-extensional interpretations to the ε-terms in later chapters.

Suppose t is a subterm of $\varepsilon x.\varphi$. Then t occurs in φ, which has lower complexity. So by inductive hypothesis, for all ϱ and $\mathcal{M}$, and so for all ϱ of form $\varrho(x/d)$ for $d \in D$, $\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \|\varphi'\|_{\mathcal{M}}^{\varrho(x/d)}$. So, for each ϱ, $\{ d \in D \mid \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \} = \{ d \in D \mid \|\varphi'\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \}$, and the result follows. □

A quick look back at the definition of ***alphabetic variants*** (Definition 5.4) should be enough to convince readers that Lemmas 2 and 3 give us

Lemma 4. *Let φ be an alphabetic variant of φ', and let t be an alphabetic variant of t'. Then for any ϱ, $\mathcal{M}$, $\|\varphi\|_{\mathcal{M}}^{\varrho} = \|\varphi'\|_{\mathcal{M}}^{\varrho}$ and $[t]_{\mathcal{M}}^{\varrho} = [t']_{\mathcal{M}}^{\varrho}$.* □

Finally, by definition, for a given x and s, t' and φ' are alphabetic variants of t and φ such that s is free for x in t' and φ'. By the definition of substitution, $t[x/s] \equiv t'[x/s]$ and $\varphi[x/s] \equiv \varphi'[x/s]$, so

$$[t[x/s]]_{\mathcal{M}}^{\varrho} = [t'[x/s]]_{\mathcal{M}}^{\varrho} \qquad \text{and} \qquad \|\varphi[x/s]\|_{\mathcal{M}}^{\varrho} = \|\varphi'[x/s]\|_{\mathcal{M}}^{\varrho}.$$

By lemma 1,

$$[t'[x/s]]_{\mathcal{M}}^{\varrho} = [t']_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} \qquad \text{and} \qquad \|\varphi'[x/s]\|_{\mathcal{M}}^{\varrho} = \|\varphi'\|_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})}.$$

Since t, t' and φ, φ' are alphabetic variants,

$$[t]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = [t']_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} \qquad \text{and} \qquad \|\varphi\|_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = \|\varphi'\|_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})}.$$

Thus,

$$[t[x/s]]_{\mathcal{M}}^{\varrho} = [t]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} \qquad \text{and} \qquad \|\varphi[x/s]\|_{\mathcal{M}}^{\varrho} = \|\varphi\|_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})}. \quad \square$$

2. A Useful Theorem

At the start of this chapter it was mentioned that the proofs in it will be referred to several times in later chapters to limit the amount of duplication. One thing that will be useful in this regard is Theorem 3, which we will be proving in this section. This theorem is the ε-calculus analogue of the fact that every Hintikka set is satisfiable (cf. e. g. [Bell and Machover, Chapter 2, §7]). Whether or not Hintikka would want to have anything to do with the sets we will deal with here, we will continue to attach his name to them.

Our definition of the appropriate sets to consider has rather a lot of clauses because all the connectives are primitives in our language. Also, of course, we need to add clauses that will enable us to handle the ε-terms, and several such clauses are required.

Definition 2. A set Ψ of formulas of $\mathcal{L}$ is an *extensional classical* ε*-Hintikka set* in $\mathcal{L}$ if the following conditions hold.

(1) If φ is atomic, then φ and $\neg\varphi$ do not both belong to Ψ.

(2) $\neg\neg\varphi \in \Psi \implies \varphi \in \Psi$

(3) $\alpha \Rightarrow \beta \in \Psi \implies \neg\alpha \in \Psi$ or $\beta \in \Psi$

(4) $\neg(\alpha \Rightarrow \beta) \in \Psi \implies$ both $\alpha \in \Psi$ and $\neg\beta \in \Psi$

(5) $\alpha \vee \beta \in \Psi \implies \alpha \in \Psi$ or $\beta \in \Psi$

(6) $\neg(\alpha \vee \beta) \in \Psi \implies$ both $\neg\alpha \in \Psi$ and $\neg\beta \in \Psi$

(7) $\alpha \wedge \beta \in \Psi \implies$ both $\alpha \in \Psi$ and $\beta \in \Psi$

(8) $\neg(\alpha \wedge \beta) \in \Psi \implies \neg\alpha \in \Psi$ or $\neg\beta \in \Psi$

(9) $\forall x.\alpha \in \Psi \implies \alpha[x/t] \in \Psi \quad$ for all $t \in$ **Terms**

(10) $\neg\forall x.\alpha \in \Psi \implies \neg\alpha[x/t] \in \Psi \quad$ for some $t \in$ **Terms**

(11) $\exists x.\alpha \in \Psi \implies \alpha[x/t] \in \Psi \quad$ for some $t \in$ **Terms**

(12) $\neg\exists x.\alpha \in \Psi \implies \neg\alpha[x/t] \in \Psi \quad$ for all $t \in$ **Terms**

(13) $\alpha[x/t] \in \Psi \quad$ for some $t \in$ **Terms** $\implies \alpha[x/\varepsilon x.\alpha] \in \Psi$

(14) for all $t \in$ **Terms**, $t = t \in \Psi$

(15) For $t_1, \ldots, t_n, s_1, \ldots, s_n \in$ **Terms**, $P \in \mathbf{Pred}_n$,

$$t_1 = s_1 \Rightarrow \ldots \Rightarrow t_n = s_n \Rightarrow Pt_1 \ldots t_n \Rightarrow Ps_1 \ldots s_n \in \Psi$$

and $\quad t_1 = s_1 \Rightarrow t_2 = s_2 \Rightarrow t_1 = t_2 \Rightarrow s_1 = s_2 \in \Psi$

(16) For all ε-terms, $\varepsilon x.\varphi = \varepsilon y.\varphi[x/y] \in \Psi$

(17) Let $t_1, \ldots, t_n, s_1, \ldots, s_n \in$ **Terms**, and write $\vec{t}$ for $\langle t_1, \ldots, t_n \rangle$, and write $\vec{s}$ for $\langle s_1, \ldots, s_n \rangle$. Suppose that $\mathbf{FV}(\varphi) \subseteq \{x_1, \ldots, x_n, y\}$ and suppose that $\mathbf{FV}(\beta) \subseteq \{x_1, \ldots, x_n, z\}$. Write $\vec{x}$ for $\langle x_1, \ldots, x_n \rangle$. Then if for all $t \in$ **Terms**

$$\varphi[\vec{x}/\vec{s}, y/t] \in \Psi \iff \beta[\vec{x}/\vec{t}, z/t] \in \Psi$$

then

$$(\varepsilon y.\varphi)[\vec{x}/\vec{s}] = (\varepsilon z.\beta)[\vec{x}/\vec{t}] \in \Psi.$$

For the remainder of this chapter we will refer to, for example, (8) of this list simply as (8), without specifying that Definition 2 is what is intended. My apologies to the reader for this lengthy definition. Note that in (17) the antecedent has a metalinguistic connective, so it is not required that a particular set of biconditional formulas actually be in Ψ. What we are about to show is that every extensional classical ε Hintikka set is satisfiable, and we will show this by constructing a structure and a valuation in which it is satisfied. Much of the baggage in this lengthy definition is required to ensure that the choice function in our structure satisfies the extensionality requirements.

The first few steps on the road to proving this are familiar.

Definition 3. Let $s, t \in$ **Terms**. Then $s \sim t \iff s = t \in \Psi$.

Lemma 5. *$\sim$ is an equivalence relation.* $\Box$

The proof is routine. Details can be found in [Bell and Machover, pp. 84–85].

Definition 4. For $t \in$ **Terms** put $[t] = \{ s \in \textbf{Terms} \mid s \sim t \}$.

Obviously, $[s] = [t] \iff s \sim t$.

Lemma 6. *Let $[t_i] = [s_i]$ for $i = 1, \ldots, n$, and let $P \in$* $\textbf{Pred}_n$. *Then we have* $Pt_1 \ldots t_n \Rightarrow Ps_1 \ldots s_n \in \Psi$. $\Box$

The proof is by application of (3) and (1) $n+1$ times to (15).

We now have enough material to construct an interpretation for $\mathcal{L}'$, the ε–free sublanguage of $\mathcal{L}$. The domain is $D = \{ [t] \mid t \in \textbf{Terms} \}$. For each $P_i \in \textbf{Pred}_n$ in $\mathcal{L}$, let

$$\mathcal{R}_i(\langle [t_1], \ldots, [t_n] \rangle) = \begin{cases} 1 & \text{if } P_i t_1 \ldots t_n \in \Psi \\ 0 & \text{otherwise} \end{cases}$$

and for each $c_j \in$ **Con**, let $\mathbf{c}(j) = [c_j]$. By construction $\mathcal{M} = \langle D, \{ R_i \mid i \in I \}, \mathbf{c} \rangle$ is a structure of the appropriate signature. Note that the cardinality of D is not greater than the cardinality of **Terms**.

It remains to define a choice function E to interpret the ε-terms. We adapt the strategy used by Leisenring (and borrowed by him from [Hermes]) for the classical extensional ε-calculus with only closed terms to the case where we allow ε-terms to be open.

Definition 5. Let $N \subseteq D$. N is *representable* by an ordered triple $\langle \varphi, y, \vec{t} \rangle$ (and the triple *represents* N) if $(\mathbf{FV}(\varphi) - \{y\}) \subseteq \{z_1, \ldots, z_n\}$, $\vec{t} = \langle t_1, \ldots, t_n \rangle$ is a sequence of terms, and (writing $\vec{z}$ for $\langle z_1, \ldots, z_n \rangle$)

$$\{ [s] \in D \mid \varphi[\vec{z}/\vec{t}, y/s] \in \Psi \} = N.$$

By (17), it follows from the fact that $\langle \varphi, x, \vec{s} \rangle$ and $\langle \beta, y, \vec{t} \rangle$ both represent N that $(\varepsilon x.\varphi)[\vec{z}/\vec{s}] \sim (\varepsilon y.\beta)[\vec{z}/\vec{t}]$. Note also that if $N \neq \emptyset$ and $\langle \varphi, x, \vec{t} \rangle$ represents N, then $[(\varepsilon x.\varphi)[\vec{z}/\vec{t}]] \in N$. This follows from (13) since $N \neq \emptyset$ implies $\varphi[\vec{z}/\vec{t}, x/t] \in \Psi$ for some term t.

So we can define our choice function $E : \mathcal{P}\, D \to D$. First we put

$$E(N) = [(\varepsilon x.\varphi)[\vec{z}/\vec{t}]]\,,$$

where $\langle \varphi, x, \vec{t} \rangle$ represents N. As we just noted, this is independent of choice of representative. If N is not representable, we can put $E(N) = d$ for an arbitrary $d \in N$. There is such a d, since $\emptyset$ is represented by $\langle x \neq x, x, \emptyset \rangle$.

So we have now defined a classical extensional ε-structure for $\mathcal{L}$, namely $\mathcal{M} = \langle D, \{ R_i \mid i \in I \}, \mathbf{c}, E \rangle$. It remains to show that there is a valuation in $\mathcal{M}$ which satisfies Ψ. Eventually we shall use the canonical valuation $\pi : x \mapsto [x]$. In ordinary predicate logic we could proceed to a straightforward inductive argument to show that $[t]^{\pi}_{\mathcal{M}} = [t]$, and that both $\varphi \in \Psi \implies \|\varphi\|^{\pi}_{\mathcal{M}} = 1$ and $\neg\varphi \in \Psi \implies \|\varphi\|^{\pi}_{\mathcal{M}} = 0$. However, things are not so easy for us. And the problem this time is not just that we need to treat the cases of terms and formulas together.

To see the difficulty, consider how we would argue for the case $\forall x.\beta$. Our inductive hypothesis would be that for all well formed expressions of complexity $\leqslant \text{comp}(\beta)$, $[t]^{\pi}_{\mathcal{M}} = [t]$, $\varphi \in \Psi \implies \|\varphi\|^{\pi}_{\mathcal{M}} = 1$, and $\neg\varphi \in \Psi \implies \|\varphi\|^{\pi}_{\mathcal{M}} = 0$. So suppose $\forall x.\beta \in \Psi$. Then for every $t \in \mathbf{Terms}$, $\beta[x/t] \in \Psi$, by (9). We would like to deduce by IH that $\|\varphi[x/t]\|^{\pi}_{\mathcal{M}} = 1$ for all $t \in \mathbf{Terms}$. However, this does not follow since there are many t sufficiently complex to make $\text{comp}(\varphi[x/t]) > n + 1$.

We will circumvent this problem by proving Lemma 7. If M is a well-formed expression with $\mathbf{FV}(M) \subseteq \{x_1, \ldots, x_n\}$

$$M \equiv M[x_1/x_1, \ldots, x_n/x_n] \equiv M[x_1/\pi(\mathbf{x_1}), \ldots, x_n/\pi(\mathbf{x_n})],$$

where $\pi(\mathbf{x_i})$ is a representative of the set $\pi(x_i)$, if we choose, as we obviously may, x_i as a representative of $\pi(x)$. Lemma 7 shows that for *arbitrary* valuations in $\mathcal{M}$ the result we want can be proved if we replace M by $M[x_1/\varrho(\mathbf{x_1}), \ldots, x_n/\varrho(\mathbf{x_n})]$, so the result we are really after will obviously follow as a special case of Lemma 7 by the syntactical fact just mentioned. But in the proof of Lemma 7 the inductive hypothesis is a claim about all valuations in $\mathcal{M}$, not a claim about all expressions of sufficiently low complexity, and $\varrho(x/[t])$ is a valuation in $\mathcal{M}$, regardless of whether $\text{comp}(\varphi[x/t]) > n + 1$.

Before stating and proving the lemma, note that for all $d \in D$, $d \neq \emptyset$. So we will choose for each $[t]$ a representative which we shall usually continue to denote denote by $\mathbf{t}$, though in some cases where confusion is impossible we will not bold, for typographical convenience (usually for cases like $[t]^{\varrho}_{\mathcal{M}}$). Also, if $\vec{x} = \langle x_1, \ldots, x_n \rangle$ is a sequence of distinct variables and ϱ is a valuation in $\mathcal{M}$, we will write $\boldsymbol{\rho}(\mathbf{x})$ for $\langle \boldsymbol{\rho}(\mathbf{x_1}), \ldots, \boldsymbol{\rho}(\mathbf{x_n}) \rangle$ the sequence of representatives of the classes to which those variables are assigned.

Lemma 7. *Let M be a well-formed expression of $\mathcal{L}$ such that $\mathbf{FV}(M) \subseteq \{x_1, \ldots, x_n\}$. Write $\vec{x}$ for $\langle x_1, \ldots, x_n \rangle$. Then for any valuation ϱ in $\mathcal{M}$,*

(1) $M \equiv t \in \mathbf{Terms} \implies [t]^{\varrho}_{\mathcal{M}} = [t[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]]$

(2) $M \equiv \varphi \in \mathbf{Wffs} \implies (\varphi[\vec{x}/\boldsymbol{\rho}(\mathbf{x})] \in \Psi \implies \|\varphi\|^{\varrho}_{\mathcal{M}} = 1)$

(3) $M \equiv \varphi \in \mathbf{Wffs} \implies (\neg\varphi[\vec{x}/\boldsymbol{\rho}(\mathbf{x})] \in \Psi \implies \|\varphi\|^{\varrho}_{\mathcal{M}} = 0)$

Proof. By induction on comp(M).

If $x \in \mathbf{Var}$, $[x]^{\varrho}_{\mathcal{M}} = \varrho(x) = [\boldsymbol{\rho}(\mathbf{x})] = [x[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]]$. If $c \in \mathbf{Con}$, $[c]^{\varrho}_{\mathcal{M}} = [c] = [c[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]]$.

For $\varphi \equiv Pt_1, \ldots, t_n$, suppose all of $t_1, \ldots, t_n$ are such that $[t_i]^{\varrho}_{\mathcal{M}} = [t_i[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]]$. Then

$$\begin{aligned}
\varphi[\vec{x}/\boldsymbol{\rho}(\mathbf{x})] \in \Psi &\implies P_i t_1[\vec{x}/\boldsymbol{\rho}(\mathbf{x})] \ldots t_n[\vec{x}/\boldsymbol{\rho}(\mathbf{x})] \in \Psi && \text{by def'n of } \mathbf{FV}(M) \\
&\implies P_i[t_1]^{\varrho}_{\mathcal{M}} \ldots [t_n]^{\varrho}_{\mathcal{M}} \in \Psi && \text{by IH} \\
&\implies \mathcal{R}_i(\langle [t_1]^{\varrho}_{\mathcal{M}}, \ldots, [t_n]^{\varrho}_{\mathcal{M}} \rangle) = 1 && \text{by def'n } \mathcal{R}_i \\
&\implies \|\varphi\|^{\varrho}_{\mathcal{M}} = 1.
\end{aligned}$$

$$
\begin{aligned}
\neg\varphi[\vec{x}/\rho(\mathbf{x})] \in \Psi &\implies \neg P_i t_1[\vec{x}/\rho(\mathbf{x})] \dots t_n[\vec{x}/\rho(\mathbf{x})] \in \Psi && \text{by def'n of } \mathbf{FV}(M) \\
&\implies \neg P_i[t_1]^{\varrho}_{\mathcal{M}} \cdots [t_n]^{\varrho}_{\mathcal{M}} \in \Psi && \text{by IH} \\
&\implies P_i[t_1]^{\varrho}_{\mathcal{M}} \cdots [t_n]^{\varrho}_{\mathcal{M}} \notin \Psi && \text{by (1)} \\
&\implies \mathcal{R}_i(\langle [t_1]^{\varrho}_{\mathcal{M}}, \dots, [t_n]^{\varrho}_{\mathcal{M}} \rangle) = 0 && \text{by def'n } \mathcal{R}_i \\
&\implies \|\varphi\|^{\varrho}_{\mathcal{M}} = 0.
\end{aligned}
$$

The case of $=$ is very similar, and the cases for the connectives are routine.

For the remaining three cases, $M \equiv \forall y.\beta$, $M \equiv \exists y.\beta$, and $M \equiv \varepsilon x.\beta$, let $\vec{z} = \langle z_1, \dots, z_m \rangle$ be a sequence of distinct variables such that $\mathbf{FV}(\beta) \subseteq \{z_1, \dots, z_m\}$. Since $\mathbf{FV}(M) \subseteq \{z_1, \dots, z_m\} - \{y\}$, the inductive hypothesis gives us, for any term t,

$$\beta[\vec{x}/\rho(\mathbf{x}), y/t] \in \Psi \implies \|\beta\|^{\varrho(y/[t]^{\varrho}_{\mathcal{M}})}_{\mathcal{M}} = 1$$

and

$$\neg\beta[\vec{x}/\rho(\mathbf{x}), y/t] \in \Psi \implies \|\beta\|^{\varrho(y/[t]^{\varrho}_{\mathcal{M}})}_{\mathcal{M}} = 0.$$

Note the following syntactical fact: If z occurs in none of the t_i in $\vec{t}$ and is free for y in β,

$$(*) \qquad \beta[y/z][\vec{x}/\vec{t}][z/t] \equiv \beta[y/z][\vec{x}/\vec{t}, z/t] \equiv \beta[\vec{x}/\vec{t}, y/t].$$

That is, we can get the same effect we get by renaming the free occurrences of y in the t_i as we did in the definition of simultaneous substitution (Definition 5.8) by *first* renaming the occurrences of y in β, then moving on to do the substitutions—and then whether the replacement of what once were the free occurrences of y

are part of the simultaneous substitution or they are left to the end is irrelevant. Now notice that $(\forall y.\beta)[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]$, if y occurs anywhere free in one of the $\boldsymbol{\rho}(\mathbf{x})$, is $\forall z.\beta[y/z][\vec{x}/\boldsymbol{\rho}(\mathbf{x})]$ where z is the first variable not in any of the $\boldsymbol{\rho}(\mathbf{x})$ but free for y in β. If y has no such free occurrence, we can say the same, taking $z \equiv y$. (Similar remarks apply, of course, if $\forall$ is replaced by $\exists$ or ε.) So

$$
\begin{aligned}
(\forall y.\beta)[\vec{x}/\boldsymbol{\rho}(\mathbf{x})] \in \Psi &\implies \forall z.\beta[y/z][\vec{x}/\boldsymbol{\rho}(\mathbf{x})] \in \Psi \\
&\implies \beta[y/z][\vec{x}/\boldsymbol{\rho}(\mathbf{x})][z/t] \in \Psi \qquad \text{for all } t \in \textbf{Terms}, \text{ by (9)} \\
&\implies \beta[\vec{x}/\boldsymbol{\rho}(\mathbf{x}), y/t] \in \Psi \qquad \text{for all } t \in \textbf{Terms}, \text{ by } (*) \\
&\implies \|\beta\|_{\mathcal{M}}^{\varrho(y/[t]_{\mathcal{M}}^{\varrho})} = 1 \qquad \text{for all } t \in \textbf{Terms}, \text{ by IH} \\
&\implies \|\forall y.\beta\|_{\mathcal{M}}^{\varrho} = 1.
\end{aligned}
$$

$$
\begin{aligned}
(\neg\forall y.\beta)[\vec{x}/\boldsymbol{\rho}(\mathbf{x})] \in \Psi &\implies \neg\beta[\vec{x}/\boldsymbol{\rho}(\mathbf{x}), y/t] \in \Psi \qquad \text{for some } t \in \textbf{Terms}, \text{ by (10) and } (*) \\
&\implies \|\beta\|_{\mathcal{M}}^{\varrho(y/[t]_{\mathcal{M}}^{\varrho})} = 0 \qquad \text{for some } t \in \textbf{Terms}, \text{ by IH} \\
&\implies \|\forall y.\beta\|_{\mathcal{M}}^{\varrho} = 0.
\end{aligned}
$$

The proofs for $\exists y.\beta$ and $\neg\exists y.\beta$ will obviously be very similar.

Finally, let $N = \{\, [t] \in D \mid \|\beta\|_{\mathcal{M}}^{\varrho(y/[t]_{\mathcal{M}}^{\varrho})} = 1 \,\}$. Then $[\varepsilon y.\beta]_{\mathcal{M}}^{\varrho} = E(N)$. But IH tells us that

$$N = \{\, [t] \in D \mid \beta[\vec{x}/\boldsymbol{\rho}(\mathbf{x}), y/t] \in \Psi \,\}$$

which is to say that $\langle \beta, y, \boldsymbol{\rho}(\mathbf{x}) \rangle$ represents N. So $[\varepsilon y.\beta]_{\mathcal{M}}^{\varrho} = [(\varepsilon y.\beta)[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]]$.

This completes the proof of Lemma 7. □

Theorem 3. *Any (classical, extensional) ε-Hintikka set is satisfiable.*

Proof. Let $\pi : \mathbf{Var} \to D$ be defined by $\pi(x) = [x]$. It follows from Lemma 7 and the fact that interpretations are independent of choice of representatives of the equivalence classes of terms that $[t]^{\pi}_{\mathcal{M}} = [t]$, that $\varphi \in \Psi \implies \|\varphi\|^{\pi}_{\mathcal{M}} = 1$, and that $\neg\varphi \in \Psi \implies \|\varphi\|^{\pi}_{\mathcal{M}} = 0$. □

3. Soundness, Completeness, and a Semantic Proof of Compactness

It is in this section that we will put to use the notion of *finitary closure operation* introduced in Chapter 2. In fact, we will have two of them on the go at once. The virtue of this approach is that it will let us give a purely semantic proof of the Compactness Theorem for extensional classical ε-structures while at the same time using essentially the same argument to prove the Completeness Theorem. We begin with

Definition 6. For a language $\mathcal{L}$, we define the *finitary semantic closure* for $\mathcal{L}$ to be the function such that for every $X \subseteq \mathbf{Wffs}$,

$$\mathbf{C}_s(X) = \{\, \varphi \in \mathbf{Wffs} \mid \text{there is a finite } Y \subseteq X \text{ such that } Y \models \varphi \,\}.$$

It is an easy exercise to show that $\mathbf{C}_s$ is a finitary closure operation on **Wffs**. Since we have not formulated our logic with a constant for falsity, we will rely on the obvious fact that $\|\forall x.x \neq x\|^{\varrho}_{\mathcal{M}} = 0$ for all ϱ and $\mathcal{M}$, and will use the abbreviation $\mathbf{f} \equiv \forall x.x \neq x$.

Lemma 8. *If* $\mathbf{f} \in \mathbf{C}_s(X)$, *then there exists a finite* $Y \subseteq X$ *which is unsatisfiable.*

Proof. By hypothesis there is a finite $Y \subseteq X$ such that $Y \models \mathbf{f}$, so there is no $\mathcal{M}$ with ϱ satisfying Y. $\square$

If $\Gamma \subseteq \mathbf{Wffs}$ and $\varphi \in \mathbf{Wffs}$, we will for the present write '$\Gamma \vdash_\varepsilon \varphi$' for "$\varphi$ is derivable from Γ in the extensional classical ε-calculus." We also want

Definition 7. For a language $\mathcal{L}$, the *deductive closure* for $\mathcal{L}$ is the function such that for every $X \subseteq \mathbf{Wffs}$,

$$\mathbf{C}_d(X) = \{\varphi \in \mathbf{Wffs} \mid X \vdash_\varepsilon \varphi\}.$$

This is also a finitary closure operation on **Wffs**. ((4) holds because we only allow finitely long derivations. For (2), the result follows by an easy inductive argument from the fact that if $\langle \varphi_1, \ldots, \varphi_n \rangle$ is a derivation of φ_n from X while $\langle \beta_1, \ldots, \beta_m \rangle$ is a derivation of β_m from $X \cup \{\varphi_n\}$, then $\langle \varphi_1, \ldots, \varphi_n, \beta_1, \ldots, \beta_m \rangle$ is a derivation of β_m from X.) We will say that a set of formulas X is *inconsistent* if $\mathbf{f} \in \mathbf{C}_d(X)$.

We will not spend much time on syntactical manipulations of the extensional classical ε-calculus. Readers interested in such things are directed to [Leisenring] or, indeed, back to [Hilbert and Bernays, vol. 2]. However, we will need a few easy to prove facts in what is to come below. First, because we have (MP), (P1) and (P2), we can use the usual proof, as found e. g. in [Bell and Machover pp. 36 37], of

Theorem 4 (Deduction Theorem). *Let* $\Gamma \subseteq \mathbf{Wffs}$, $\varphi, \beta \in \mathbf{Wffs}$. *Then*

$$\Gamma \cup \{\varphi\} \vdash_\varepsilon \beta \iff \Gamma \vdash_\varepsilon \varphi \Rightarrow \beta. \quad \square$$

We will also want the following series of Lemmas.

Lemma 9(a). (1) $\mathbf{f} \in \mathbf{C}_s(X) \iff \mathbf{C}_s(X) = \mathbf{Wffs}$

(2) $\varphi \in \mathbf{C}_s(X) \iff \mathbf{f} \in \mathbf{C}_s(X \cup \{\neg\varphi\})$

(3) $\neg\varphi \in \mathbf{C}_s(X) \iff \mathbf{f} \in \mathbf{C}_s(X \cup \{\varphi\})$

(4) $Max_{\mathbf{C}_s}(X) \implies \mathbf{C}_s(X) = X$

Proof. (1) For ($\Longrightarrow$), if $Z \models \mathbf{f}$, then Z is unsatisfiable, so $Z \models \varphi$ for all $\varphi \in \mathbf{Wffs}$. The converse is obvious.

(2) For ($\Longrightarrow$), there is a finite $Y \subseteq X$ such that $Y \models \varphi$. Thus the finite set $Y \cup \{\neg\varphi\}$ is unsatisfiable, and $Y \cup \{\neg\varphi\} \models \mathbf{f}$. But $Y \cup \{\neg\varphi\} \subseteq X \cup \{\neg\varphi\}$. For ($\Longleftarrow$), we know by lemma 8 that there is a finite $Y \subseteq X \cup \{\neg\varphi\}$ which is unsatisfiable. But $Y - \{\neg\varphi\} \models \varphi$, since if $\mathcal{M}, \varrho \models Y - \{\neg\varphi\}$, $\|\varphi\|^{\varrho}_{\mathcal{M}} \neq 0$, or else Y would be satisfiable.

(3) follows from (2) and the fact that $\{\neg\neg\varphi\} \models \varphi$.

(4) Since $Max_{\mathbf{C}_s}(X)$, $\mathbf{f} \notin \mathbf{C}_s(X)$, and so is not in $\mathbf{C}_s(\mathbf{C}_s(X))$ by (2) of the definition of finitary closure, and $X \subseteq \mathbf{C}_s(X)$, by (1) of that definition, the result follows by the definition of maximality. □

Lemma 9(b). (1) $\mathbf{f} \in \mathbf{C}_d(X) \iff \mathbf{C}_d(X) = \mathbf{Wffs}$

(2) $\varphi \in \mathbf{C}_d(X) \iff \mathbf{f} \in \mathbf{C}_d(X \cup \{\neg\varphi\})$

(3) $\neg\varphi \in \mathbf{C}_d(X) \iff \mathbf{f} \in \mathbf{C}_d(X \cup \{\varphi\})$

(4) $Max_{\mathbf{C}_d}(X) \implies \mathbf{C}_d(X) = X$

Proof. (1) Suppose $X \vdash_e \forall x.(x \neq x)$. For any $t \in \mathbf{Terms}$, $\forall x.(x \neq x) \Rightarrow t \neq t$ is an axiom, so $X \vdash_e t \neq t$ by (MP). But $t = t$ is an axiom, so the result follows by propositional logic. The converse is obvious.

(2) If $X \cup \{\neg\varphi\}$ is inconsistent, $X \cup \{\neg\varphi\} \vdash_e \varphi$ by (1). So $X \vdash_e \neg\varphi \Rightarrow \varphi$ by

the Deduction Theorem. But $(\neg\varphi \Rightarrow \varphi) \Rightarrow \varphi$ is a propositional tautology, so $\vdash_\varepsilon (\neg\varphi \Rightarrow \varphi) \Rightarrow \varphi$, so $X \vdash_\varepsilon \varphi$ by (MP). The converse is obvious.

(3) Similar to (2).

(4) The same as (4) of Lemma 9(a). □

Lemma 10(a). *Let $\mathcal{M}$ be a structure for $\mathcal{L}$ and ϱ a valuation in $\mathcal{M}$.*

(1) $\|\varphi \Leftrightarrow \beta\|^\varrho_\mathcal{M} = 1 \iff \|\varphi\|^\varrho_\mathcal{M} = \|\beta\|^\varrho_\mathcal{M}$

(2) $\|\forall x.(\varphi \Leftrightarrow \beta)\|^\varrho_\mathcal{M} = 1 \implies [\varepsilon x.\varphi]^\varrho_\mathcal{M} = [\varepsilon x.\beta]^\varrho_\mathcal{M}$

Proofs. (1) follows from the interpretations of $\wedge$ and $\Rightarrow$ and the definition of $\Leftrightarrow$. For (2),

$$\begin{aligned}
\|\forall x.(\varphi \Leftrightarrow \beta)\|^\varrho_\mathcal{M} = 1 &\iff \|\varphi \Leftrightarrow \beta\|^{\varrho(x/d)}_\mathcal{M} = 1 \qquad \text{for all } d \in D \\
&\iff \|\varphi\|^{\varrho(x/d)}_\mathcal{M} = \|\beta\|^{\varrho(x/d)}_\mathcal{M} \qquad \text{for all } d \in D \\
&\iff \{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_\mathcal{M} = 1 \,\} = \{\, d \in D \mid \|\beta\|^{\varrho(x/d)}_\mathcal{M} = 1 \,\} \\
&\implies [\varepsilon x.\varphi]^\varrho_\mathcal{M} = [\varepsilon x.\beta]^\varrho_\mathcal{M} \quad \square
\end{aligned}$$

We state, for convenience, the following obvious fact in the form of a lemma.

Lemma 10(b). $\Gamma \vdash_\varepsilon \forall x.(\varphi \Leftrightarrow \beta) \implies \Gamma \vdash_\varepsilon \varepsilon x.\varphi = \varepsilon x.\beta$. □

Lemma 11(a). (1) $\|\neg\forall x.\varphi\|^\varrho_\mathcal{M} = 1 \implies \|\neg\varphi[x/\varepsilon x.\neg\varphi]\|^\varrho_\mathcal{M} = 1$

(2) $\|\exists x.\varphi\|^\varrho_\mathcal{M} = 1 \implies \|\varphi[x/\varepsilon x.\varphi]\|^\varrho_\mathcal{M} = 1$

Proof. For (1),

$$\|\neg\forall x.\varphi\|^{\varrho}_{\mathcal{M}} = 1 \implies \|\forall x.\varphi\|^{\varrho}_{\mathcal{M}} = 0$$

$$\implies \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 0 \qquad \text{for some } d \in D$$

$$\implies \|\neg\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 1 \qquad \text{for some } d \in D.$$

So put $N = \{\, d \in D \mid \|\neg\varphi\|^{\varrho(x/d)}_{\mathcal{M}}\} = 1$. Since $N \neq \emptyset$, $E(N) \in N$, so $[\varepsilon x.\neg\varphi]^{\varrho}_{\mathcal{M}} \in N$. Thus $\|\neg\varphi\|^{\varrho(x/[\varepsilon x.\neg\varphi]^{\varrho}_{\mathcal{M}})}_{\mathcal{M}} = 1$, and so $\|\neg\varphi[x/\varepsilon x.\neg\varphi]\|^{\varrho}_{\mathcal{M}} = 1$, by Theorem 2.

The proof for (2) is obviously very similar. □

Lemma 11(b). (1) $\Gamma \vdash_{\varepsilon} \neg\forall x.\varphi \implies \Gamma \vdash_{\varepsilon} \neg\varphi[x/\varepsilon x.\neg\varphi]$

(2) $\Gamma \vdash_{\varepsilon} \exists x.\varphi \implies \Gamma \vdash_{\varepsilon} \varphi[x/\varepsilon x.\varphi]$.

Proof. (2) is just an application of (MP) to (ε). (1) follows from (ε) and the fact that in classical first order logic $\neg\forall x.\varphi \Rightarrow \exists x.\neg\varphi$ is derivable. □

Finally we are ready to begin proving the basic theorems for our syntax and semantics.

Theorem 5 (Soundness Theorem). *Let* $\Gamma \subseteq$ **Wffs** *and* $\varphi \in$ **Wffs**. *Then*

$$\Gamma \vdash_{\varepsilon} \varphi \implies \Gamma \models \varphi.$$

Proof. By the Soundness Theorem for the classical predicate calculus we know that the axioms (P1)-(P11) and (Q1)-(Q6) are valid in, and (MP) is truth preserving in, any structure for $\mathcal{L}$, and so in particular for any classical extensional ε-structure for $\mathcal{L}$. So it suffices to show that (α), (ε), and (Ack) are valid in any classical extensional ε-structure.

For (Ack) we have by Lemma 10(a)(2) that

$$\|\forall x.(\varphi \Leftrightarrow \beta)\|^{\varrho}_{\mathcal{M}} = 1 \implies [\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = [\varepsilon x.\beta]^{\varrho}_{\mathcal{M}},$$

so by our definition of interpretation

$$\|\forall x.(\varphi \Leftrightarrow \beta)\|^{\varrho}_{\mathcal{M}} = 1 \implies \|\varepsilon x.\varphi = \varepsilon x.\beta\|^{\varrho}_{\mathcal{M}} = 1,$$

and so, since every formula is assigned a value in **2** in each $\mathcal{M}$ and ϱ,

$$\|\forall x.(\varphi \Leftrightarrow \beta) \Rightarrow \varepsilon x.\varphi = \varepsilon x.\beta\|^{\varrho}_{\mathcal{M}} = 1.$$

The validity of (ε) follows similarly by Lemma 11(a). The result is obvious for (α), by the definition of substitution and lemma 4. □

Corollary. *For any* $X \subseteq \mathbf{Wffs}$, $\mathbf{C}_d(X) \subseteq \mathbf{C}_s(X)$. □

The following theorem is a trivial consequence of Theorem 2.3.

Theorem 6. (1) *If* $\mathbf{f} \notin \mathbf{C}_s(X)$, *then there is a* Y *such that* $X \subseteq Y$ *and* $Max_{\mathbf{C}_s}(Y)$.

(2) *If* $\mathbf{f} \notin \mathbf{C}_d(X)$, *then there is a* Y *such that* $X \subseteq Y$ *and* $Max_{\mathbf{C}_d}(Y)$. □

For each X, these two maximal sets are in fact the same, and they are close kin of Henkin's maximal consistent sets. We set about proving this now.

Theorem 7. (1) $Max_{\mathbf{C}_s}(X) \implies (\varphi \in X \iff \neg\varphi \notin X)$.

(2) $Max_{\mathbf{C}_d}(X) \implies (\varphi \in X \iff \neg\varphi \notin X)$.

Proof. (1) ($\Longrightarrow$) If $\neg\varphi \in X$, then $\{\varphi, \neg\varphi\}$ is an unsatisfiable finite subset of X, so $\mathbf{C}_s(X) = X = \mathbf{Wffs}$, so X is not maximal. ($\Longleftarrow$) By lemma 9(a)(1), $\neg\varphi \notin X =$

$\mathbf{C}_s(X) \iff \mathbf{f} \notin \mathbf{C}_s(X \cup \{\neg\neg\varphi\})$. But $\mathbf{C}_s(X) \subseteq \mathbf{C}_s(X \cup \{\neg\neg\varphi\})$, since $\mathbf{C}_s$ is a finitary closure operation, so $\mathbf{C}_s(X \cup \{\neg\neg\varphi\}) = X$. Since $\{\neg\neg\varphi\} \models \varphi$, $\varphi \in X$.

(2) The proofs are the same, with the obvious modifications. That is, we rely on the fact of propositional logic that every formula is derivable from $\{\varphi, \neg\varphi\}$ to prove the first half, use lemma 9(b) rather than 9(a) and the fact that $\neg\neg\varphi \Rightarrow \varphi$ is a propositional tautology in place of the fact that $\{\neg\neg\varphi\} \models \varphi$ to prove the second half. □

It is very straightforward to prove the following using the definition of interpretation and classical propositional logic.

Lemma 12. *For $i \in \{d, s\}$, if $X \subseteq$ **Wffs**, $\alpha, \beta \in$ **Wffs**, then*

(1) $\neg\neg\alpha \in \mathbf{C}_i(X) \iff \alpha \in \mathbf{C}_i(X)$

(2) $\alpha \wedge \beta \in \mathbf{C}_i(X) \iff \alpha \in \mathbf{C}_i(X)$ and $\beta \in \mathbf{C}_i(X)$

(3) $\alpha \vee \beta \in \mathbf{C}_i(X) \iff \alpha \in \mathbf{C}_i(X)$ or $\beta \in \mathbf{C}_i(X)$

(4) $\alpha \Rightarrow \beta \in \mathbf{C}_i(X) \iff \neg\alpha \in \mathbf{C}_i(X)$ or $\beta \in \mathbf{C}_i(X)$. □

Only slightly more difficult to prove is

Lemma 13. *For $i \in \{s, d\}$, if $Max_{\mathbf{C}_i}(X)$, then*

(1) $\forall x.\varphi \in X \iff \varphi[x/t] \in X$ *for all* $x \in$ **Terms**

(2) $\neg\exists x.\varphi \in X \iff \neg\varphi[x/t] \in X$ *for all* $x \in$ **Terms**

(3) $\neg\forall x.\varphi \in X \iff \neg\varphi[x/t] \in X$ *for some* $x \in$ **Terms**

(4) $\exists x.\varphi \in X \iff \varphi[x/t] \in X$ *for some* $x \in$ **Terms**

Proof. First, $\{\forall x.\varphi\} \models \varphi[x/t]$ and $\forall x.\varphi \Rightarrow \varphi[x/t]$ is an axiom for each $t \in$

Terms. Since $\mathbf{C}_i(X) = X$ by lemma 9, it follows that $\forall x.\varphi \in X \implies \varphi[x/t] \in X$ for all $t \in$ **Terms**. Similarly, $\neg\exists x.\varphi \in X \implies \neg\varphi[x/t]$ for all $t \in$ **Terms**.

Now suppose $\neg\forall x.\varphi \in X$. By Lemma 11(a)(1) we get $\{\neg\forall x.\varphi\} \models \neg\varphi[x/\varepsilon x.\neg\varphi]$, and so $\neg\varphi[x/\varepsilon x.\neg\varphi] \in \mathbf{C}_s(X) = X$, while the result for $i = d$ follows from Lemma 11(b)(1). The result for $\exists x.\varphi$ follows similarly from Lemma 11(2).

Now suppose $\varphi[x/t] \in X$ for all $t \in$ **Terms**. If $\forall x.\varphi \notin X$, $\neg\forall x.\varphi \in X$ by Theorem 7. But then $\neg\varphi[x/t] \in X$ for some $t \in$ **Terms**, as was just shown, which contradicts Theorem 7. We get the other converses similarly. □

It is now fairly straightforward to prove

Theorem 8. *For $i \in \{d, s\}$, $Max_{\mathbf{C}_i}(X) \implies X$ is a classical extensional ε Hintikka set.*

Proof. That (1) holds obviously follows from Theorem 7. The cases (2)–(12) follow easily from Lemmas 12 and 13.

Let $i = s$. For (13), $\{\alpha[x/t]\} \models \exists x.\alpha$, so $\{\alpha[x/t]\} \models \alpha[x/\varepsilon x.\alpha]$ by lemma 11(a). (14) and (15) are valid, and $\emptyset$ is a finite subset of X. For $i = d$, (13) follows from axioms (Q6) and (ε). (14) and (15) are axioms.

For (16), the case of $i = d$ is trivial because of the (α) axioms. For $i = s$, it is enough to show that

$$\{\, d \in D \mid \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \,\} = \{\, d \in D \mid \|\varphi[x/y]\|_{\mathcal{M}}^{\varrho(y/d)} = 1 \,\}.$$

This is obvious if y is free for x in φ. If y is not free for x in φ, the result follows by the definition of substitution and lemma 4.

Finally, for (17) suppose that for all $t \in$ **Terms**,

$$\varphi[\vec{x}/\vec{s}, y/t] \in X \iff \beta[\vec{x}/\vec{t}, z/t] \in X.$$

Choose a $w \in$ **Var** which does not occur in φ, in β, or in any of the s_i or t_i. Consider $\varphi[\vec{x}/\vec{s}, y/w]$ and $\beta[\vec{x}/\vec{t}, z/w]$. The free occurrences of y in φ and of z in β are precisely the places where w occurs in these new formulas. So

$$\varphi[\vec{x}/\vec{s}, y/w][w/t] \in X \iff \varphi[\vec{x}/\vec{s}, y/t] \in X$$
$$\iff \beta[\vec{x}/\vec{t}, z/t] \in X \iff \beta[\vec{x}/\vec{t}, z/w][w/t] \in X.$$

By Lemma 12(2) and (4) and the definition of $\Leftrightarrow$, it follows that $\varphi[\vec{x}/\vec{s}, y/w][w/t] \Leftrightarrow \beta[\vec{x}/\vec{t}, z/w][w/t] \in X$, for all $t \in$ **Terms**. So $\forall w. (\varphi[\vec{x}/\vec{s}, y/w] \Leftrightarrow \beta[\vec{x}/\vec{t}, z/w]) \in X$, by Lemma 13(1). Now $\varepsilon w.(\varphi[\vec{x}/\vec{s}, y/w]) = \varepsilon w.(\beta[\vec{x}/\vec{t}, z/w]) \in X$, by Lemmas 10(a) and (b). But it is not hard to see that $\varepsilon w.(\varphi[\vec{x}/\vec{s}, y/w]) = (\varepsilon y.\varphi)[\vec{x}/\vec{t}] \in X$, by (16), since w is just holding the place of y in φ. Similar reasoning works for β, so the result follows by (15). □

Connoisseurs of such things will notice that it is in its handling of (9) that the advantage of the ε-structures over ordinary first order structures can be felt. A maximal consistent set for the ε-calculus will have its witnessing constants built in, so we need not pause to add them as is done in Henkin completeness proofs for the ordinary first order predicate calculus.

We now have all the elements we need to give a proof of

Theorem 9 (Compactness Theorem). *If $X \subseteq$ **Wffs** is such that every finite* *subset of X is satisfiable, then X is satisfiable in some $\mathcal{M}$ such that $card(M) \leqslant card(\mathcal{L})$.*

Proof. By the converse of lemma 8, $\mathbf{f} \notin X$. So X is contained in a set Y maximal under the finitary semantic closure by Theorem 6, hence in a ε Hintikka set, by Theorem 8. But then it is satisfiable by Theorem 3. □

Note that we have just given a purely semantic proof of the Compactness Theorem. That is, we can carry out the proof of this theorem without any mention of the axioms and rules of inference of the calculus as we have presented it. All the mention of these things above result from our desire not to run through what is essentially the same argument a second time to prove

Theorem 10 (Completeness Theorem). *Let* $\Gamma \subseteq \mathbf{Wffs}$, $\varphi \in \mathbf{Wffs}$.

(1) *If Γ is consistent, then Γ is satisfiable in some model such that $card(\mathcal{M}) \leqslant card(\mathcal{L})$.*

(2) $\Gamma \models \varphi \Rightarrow \Gamma \vdash_{\varepsilon} \varphi$

Proof. (1) is immediate from theorems 6, 8, and 3. For (2), suppose $\Gamma \models \varphi$. Then $\Gamma \cup \{\neg\varphi\}$ is unsatisfiable. So by (1) $\Gamma \cup \{\neg\varphi\}$ is inconsistent, and so by Lemma 9(b)(1), $\Gamma \vdash_{\varepsilon} \varphi$. □

4. ε–Theorems and Other Theorems for ε

It would be possible at this point to prove that many of the standard model theoretical results for standard classical logic also hold for the classical extensional ε-calculus. Rather than doing so here, I will rest content with referring readers to the end of chapter 1 of [Leisenring], and especially to [da Costa 1980], where many model theoretic results are proved for this semantics. These include the Löwenheim Skolem Theorem, the Craig Interpolation Lemma, various definability theorems, including the Beth Definability Theorem, and several others. Here we shall content ourselves with proving some of the best known results about the ε–calculus.

The best known results of the classical ε–calculus are Hilbert's ε–theorems, which are essentially proofs that the ε-axioms and ε-terms are eliminable from derivations starting and ending with ε-free formulas. The original proof of these theorems takes up more than 50 pages of [Hilbert–Bernays, vol. 2], because it is a syntactical proof and so spells out just how to go about eliminating occurrences of these terms and axioms. Asser and Rasiowa in the 1950's both gave model theoretic proofs of the ε-axioms, and Asser claims it as one of the virtues of his systematic treatment of the ε-calculus that "on the basis of the above considerations, ... they follow almost trivially" [Asser, p.63, my translation]. On the basis of finitist scruples Leisenring returns to giving a syntactical proof which, while more efficient than Hilbert and Bernays' original, still occupies 16 pages. Those who share Leisenring's scruples are referred to his proof, which can be found in [Leisenring pp. 64–80].

Theorem 11 (Second ε–Theorem). *Let $X \subseteq$ **Wffs**, $\varphi \in$ **Wffs**, and assume that all members of $X \cup \{\varphi\}$ are ε-free. Then*

$$X \vdash_{\varepsilon} \varphi \implies X \vdash_{PC} \varphi.$$

Proof. By the Soundness Theorem for the ε-calculus, $X \vdash_{\varepsilon} \varphi \implies X \models \varphi$.

But every classical extensional ε-structure is an ordinary structure and the interpretations of the members of X and φ are independent of E in these structures. Conversely, every ordinary structure gives rise to ε structures. So $X \models_{PC} \varphi$. But then $X \vdash_{PC} \varphi$ by the Completeness Theorem for classical first order logic. □

If we write $\vdash_{EC}$ for derivability in the classical propositional calculus, we can prove similarly the

Theorem 12 (First ε–Theorem). *Let $X \subseteq$ **Wffs**, $\varphi \in$ **Wffs**, and assume that all members of $X \cup \{\varphi\}$ are ε-free and quantifier free. Then*

$$X \vdash_\varepsilon \varphi \implies X \vdash_{EC} \varphi.$$

Proof. Obviously, by the Second ε-theorem it is enough to show $X \vdash_{PC} \varphi \implies X \vdash_{EC} \varphi$. But $X \vdash_{PC} \varphi \implies X \models_{PC} \varphi$ by the Soundness Theorem for the classical predicate calculus. But since the members of X and φ are quantifier free, for any valuation in any structure their truth-values are independent of the quantifier rules and so (in non-atomic cases) depend only on **2**-valued truth-functional semantics. So by the well-known completeness of this semantics for the classical propositional calculus, $X \vdash_{EC} \varphi$. □

In closing this section, we will give two applications of the Second ε-Theorem. To state these results in as convenient and familiar a form as possible, we will presume that the syntax of the ε-calculus has been modified to allow function symbols. This adds complexity to the earlier definitions and proofs, but creates no difficulties of principle. As usual, a structure assigns a function $D^n \to D$ to an n place function symbol.

It will be convenient to list the following facts, which now should be obvious by the Completeness Theorem and Definition 1 (for (3), take the converse of Lemma 11(a).)

Lemma 14.

(1) $X \vdash_\varepsilon \forall x.\varphi \implies X \vdash_\varepsilon \varphi[x/t]$ *for all* $t \in$ **Terms**.

(2) $X \vdash_\varepsilon \varphi[x/t]$ *for some* $t \in$ **Terms** $\implies X \vdash_\varepsilon \exists x.\varphi$.

(3) $X \vdash_\varepsilon \varphi[x/\varepsilon x.\neg\varphi] \implies X \vdash_\varepsilon \forall x.\varphi$.

(4) $X \vdash_\varepsilon \exists x.\varphi \implies X \vdash_\varepsilon \varphi[x/\varepsilon x.\varphi]$. □

For the remainder of this chapter we will refer to, for example, Lemma 14(3) simply as (3).

Lemma 15. *Let* $X \subseteq$ **Wffs**, $\varphi \in$ **Wffs** *where* $\varphi \equiv \exists x_1 \ldots \exists x_n \forall y.\beta$ *where* $n \geqslant 0$, *the variables* $x_1, \ldots, x_n, y$ *are all distinct. Let* $\mathcal{L}^+$ *be the language which results by adding a new* n*-place function symbol* g *to the vocabulary of* $\mathcal{L}$. *Write* $\vdash_\mathcal{L}, \vdash_{\mathcal{L}^+}$ *for derivability in the* ε*-calculus based on the respective languages. Then*

$$X \vdash_\mathcal{L} \exists x_1 \ldots \exists x_n \forall y.\beta \iff X \vdash_{\mathcal{L}^+} \exists x_1 \ldots \exists x_n.\beta[y/g(x_1, \ldots, x_n)].$$

Proof. To keep things simple, we will assume $n = 1$. The proof generalizes easily. For ($\Longrightarrow$), note that any proof in $\mathcal{L}$ is also a proof in $\mathcal{L}^+$. So

$$\begin{aligned} X \vdash_\mathcal{L} \exists x.\forall y.\beta &\implies X \vdash_{\mathcal{L}^+} \exists x.\forall y.\beta \\ &\implies X \vdash_{\mathcal{L}^+} \forall y.\beta[x/\varepsilon x.(\forall y.\beta)] && \text{by (4)} \\ &\implies X \vdash_{\mathcal{L}^+} \beta[x/\varepsilon x.(\forall y.\beta)][y/g(\varepsilon x.(\forall y.\beta))] && \text{by (1).} \end{aligned}$$

But $\beta[x/\varepsilon x.(\forall y.\beta)][y/g(\varepsilon x.(\forall y.\beta))] \equiv \beta[y/g(x)][x/\varepsilon x.(\forall y.\beta)]$, and so we have by (2) that $X \vdash_{\mathcal{L}^+} \exists x.\beta[y/g(x)]$.

For ($\Longleftarrow$), let $\langle \beta_1, \ldots, \beta_m \rangle$ be a derivation of $\exists x.\beta[y/g(x)]$ from X in $\mathcal{L}^+$. Take $\varepsilon y.\neg\beta$. Rename all the bound variables in $\varepsilon y.\neg\beta$ by variables which occur nowhere in the derivation. C[illegible] the result $\varepsilon w.\beta'$.

Now, for $i = 1, \ldots n$, let β_i' be β_i with all occurrences of $g(t)$ replaced by $\varepsilon w.\beta'[x/t]$ (t is free for x in $\varepsilon w.\beta'$ by our relabelling). We show that $X \vdash_{\mathcal{L}} \beta_i'$ for $i = 1, \ldots n$.

Since X is a set of formulas of $\mathcal{L}$, if $\beta_i \in X$, g does not occur in X, and so $\beta_i \equiv \beta_i'$. If β_i is an axiom of $\mathcal{L}'$, it is not hard to see that the substitutions involved in moving from β_i to β_i' will leave us with a formula of the same form, hence with an axiom of $\mathcal{L}$. Finally, if $\beta_i \equiv \beta_j \Rightarrow \beta_l$ for $1 \leqslant i, j \leqslant l \leqslant n$, then β_l' follows from β_j' and β_i' by (MP).

Thus $X \vdash_{\mathcal{L}} \beta_n'$, i. e. $X \vdash_{\mathcal{L}} \exists x.\beta[y/\varepsilon w.\neg\beta']$. Soundness, Lemma 4 and Completeness tell us that $X \vdash_{\mathcal{L}} \exists x.\beta[y/\varepsilon y.\neg\beta]$.

By the restrictions we have placed on the variables in our statement of the lemma, for any term s, $\beta[y/\varepsilon y.\neg\beta][x/s] \equiv \beta[x/s][y/\varepsilon y.\neg\beta[x/s]]$. So

$$X \vdash_{\mathcal{L}} \exists x.\beta[y/\varepsilon y.\neg\beta]$$

$$X \vdash_{\mathcal{L}} \beta[y/\varepsilon y.\neg\beta][x/\varepsilon x.\beta[y/\varepsilon y.\neg\beta]] \qquad \text{by (4)}$$

$$X \vdash_{\mathcal{L}} \beta[x/\varepsilon x.\beta[y/\varepsilon y.\neg\beta]][y/\varepsilon y.\neg\beta[x/\varepsilon x.\beta[y/\varepsilon y.\neg\beta]]]$$

$$X \vdash_{\mathcal{L}} \forall y.\beta[x/\varepsilon x.\beta[y/\varepsilon y.\neg\beta]] \qquad \text{by (3)}$$

$$X \vdash_{\mathcal{L}} \exists x.\forall y.\beta \qquad \text{by (2)} \quad \square$$

Lemma 16. *Let $X \subseteq$ **Wffs**, $\varphi \in$ **Wffs** where $\varphi \equiv \forall x_1 \ldots \forall x_n \exists y.\beta$ where $n \geqslant 0$, the variables $x_1, \ldots, x_n, y$ are all distinct. Let $\mathcal{L}^+$ be the language which results by adding a new n-place function symbol g to the vocabulary of $\mathcal{L}$. Then*

$$\forall x_1 \ldots \forall x_n \exists y.\beta \text{ is satisfiable} \iff \forall x_1 \ldots \forall x_n.\beta[y/g(x_1, \ldots, x_n)] \text{ is satisfiable.}$$

Proof. ($\Longleftarrow$) Any $\mathcal{M}$ and ϱ which make $\forall x_1 \ldots \forall x_n.\beta[y/g(x_1, \ldots, x_n)]$ true are such that for every $d_1, \ldots, d_n \in D$ there is a $d \in D$ (viz. $[g(x_1, \ldots, x_n)]^{\varrho}_{\mathcal{M}}$) such that

$$\|\beta\|^{\varrho(x_1/d_1)\ldots(x_n/d_n)(y/d)}_{\mathcal{M}} = 1.$$

But then $\|\forall x_1 \ldots \forall x_n \exists y.\beta\|^{\varrho}_{\mathcal{M}} = 1$.

For ($\Longrightarrow$), suppose $\|\forall x_1 \ldots \forall x_n \exists y.\beta\|^{\varrho}_{\mathcal{M}} = 1$. Then we can extend our structure $\mathcal{M}$ for a structure for $\mathcal{L}^+$ by interpreting

$$[g(x_1, \ldots, x_n)]^{\varrho}_{\mathcal{M}} = [\varepsilon y.\beta]^{\varrho}_{\mathcal{M}}$$

and the result is obvious. □

Definition 8. A formula is *prenex* if it has the form $Q_1 x_1 \ldots Q_n x_n.\beta$ where β is quantifier free and where $n \geqslant 0$, and for each i, Q_i is $\forall$ or $\exists$. A prenex formula is in *normal form* if the x_i are all distinct, and all of them occur in β.

Let φ be a prenex formula in normal form. We outline a procedure for getting an $\exists$-free formula from φ. Take the leftmost $\exists$-quantifier Q_r. There are two cases to consider. (1) If no $\forall$ occurs before Q_r, choose a new (i. e. not in the language, not used at an earlier stage of this process) constant c, replace x_r by c in β, and delete $Q_r x_r$ from the prefix. (2) If $\forall x_1 \ldots \forall x_m$ occurs before Q_r, $m > 0$, choose a new m-place function symbol f. Replace x_r in β by $f(x_1, \ldots, x_m)$, and delete $Q_r x_r$ from the prefix. Repeat this process until what is left is an $\exists$ free formula. The new constants and function symbols are called the ***Skolem constants*** and the ***Skolem functions***, and the resulting formula is called the ***Skolem resolution*** of φ.

If we rewrite the preceding paragraph, replacing the references to $\exists$ by $\forall$ and vice-versa, we can define the ***Herbrand constants*** and ***Herbrand functions*** and the ***Herbrand resolution*** of φ.

It is a straightforward matter to prove that for any ε-free formula φ there is a logically equivalent formula φ' which is in prenex normal form. Also, it is an easy matter to give inductive arguments on the number of $\forall$-quantifiers and $\exists$ quantifiers respectively, using lemmas 15 and 16, to show

Lemma 17. *Let φ be a formula of $\mathcal{L}$ in prenex normal form, X a set of formulas of $\mathcal{L}$, $\mathcal{L}^+$ the language which results by adding the Herbrand constants and Herbrand functions for φ to the vocabulary of $\mathcal{L}$, and φ' the Herbrand resolution of φ. Then*

$$X \vdash_{\mathcal{L}} \varphi \iff X \vdash_{\mathcal{L}^+} \varphi'. \quad \Box$$

Lemma 18. *Let φ be a formula of $\mathcal{L}$ in prenex normal form, X a set of formulas of $\mathcal{L}$, $\mathcal{L}^+$ the language which results by adding the Skolem constants and Skolem functions for φ to the vocabulary of $\mathcal{L}$, and φ' the Skolem resolution of φ. Then*

$$\varphi \text{ is satisfiable} \iff \varphi' \text{ is satisfiable}. \quad \Box$$

Since each ϵ free formula is equivalent to a prenex formula in normal form, we can apply Lemma 17 and the Second ϵ Theorem to prove

Theorem 13. *If $X \cup \{\varphi\}$ is a set of formulas of the first-order predicate calculus, then there is an $\forall$-free formula of the first-order predicate calculus φ' such that $X \vdash_{PC} \varphi \iff X \vdash_{PC} \varphi'$.* □

We conclude this chapter by noting that Lemma 18 gives us

Theorem 14 (Satisfiability Theorem). *For every formula φ of the classical first order predicate calculus there is an $\exists$-free formula φ' such that φ is satisfiable if and only if φ' is satisfiable.* □

CHAPTER VIII

SEMANTICS FOR CLASSICAL HILBERTIAN ε—CALCULUS AND CLASSICAL ε—CALCULUS

Readers who ploughed through Chapter 5 will recall that the classical Hilbertian ε-calculus is what results when we drop the (Ack) axiom schema from the calculus dealt with in Chapter 7 and replace it with

(A6) $$t_1 = t_2 \implies \varepsilon x.\varphi[y/t_1] = \varepsilon x.\varphi[y/t_2];$$

and the classical ε-calculus is what results if we change (Q5) and (Q6) to (Q5′) and (Q6′), and drop (Ack) without replacing it.

[Asser] presented a semantics for the classical Hilbertian ε-calculus. We will give essentially the same semantics for it in this chapter, but will describe it in terms of the modifications it requires to the semantics in Chapter 7. We will then give a semantics for the classical ε-calculus which requires modifications in the same spirit. While the semantics of these calculi is of some intrinsic interest, the modifications we need to make here will also be useful in Chapter 12 where we will use them as models for a more general method of designing structures to interpret calculi satisfying different sorts of extensionality conditions.

1. Semantics for the Classical Hilbertian ε–Calculus

If our goal is to construct an interpretation of the classical Hilbertian ε-calculus, it is not hard to see what modifications need to be made to the choice function

E used in Chapter 7 to interpret the ε-terms. We can no longer interpret $\varepsilon x.\varphi$ (under ϱ) by merely choosing an element out of the subset of the domain of our interpretation the members of which when "plugged in" for x in φ satisfy φ (under ϱ). For ($A6$) leaves open the possibility that two predicates have the same such subset but non-equal ε-terms. Asser solved this problem by making the choice function E depend not only on the just described subset, but also on the syntax of the ε-term to be interpreted. We will present the details of how he did this, modified to conform to our notation and terminology. This will make fairly obvious the modifications we will need to make to the proofs in Chapter 7 in order to prove corresponding results for the Hilbertian case. This will allow us to give brief proofs of the theorems by referring extensively to Chapter 7.

Definition 1 (Ground terms). We will define, for each ε–term t of $\mathcal{L}$, a ground term for t. Before proceeding, we remind the reader that an occurrence of a term t is free in a well-formed expression M if there is no variable x such that $x \in \mathbf{FV}(t)$ and that instance of t occurs in some subterm or subformula N of M such that $x \notin \mathbf{FV}(N)$, and that the set $\mathbf{FT}(M)$ of free terms of M is the set of terms which have a *free occurrence in M which is not a proper subterm of a free occurrence of a term in M*. (We might say that only the ***biggest***, or the ***highest level of***, free occurrences of terms get into that set.) So, possibly, $\mathbf{FV}(t) \not\subseteq \mathbf{FT}(t)$.

We begin by numbering the bound variables and the free terms of an ε–term t as follows.

(1) Set $j = 1$.

(2) Starting at the left, find the first occurrence of a string of form $\forall x$, $\exists x$, or εx such that x has not yet been numbered, or of an occurrence of a proper subterm s of t such that $s \in \mathbf{FT}(t)$ and s has not yet been numbered. If you find such an s, assign j to it and go to (3). If you find $\forall x$, $\exists x$, or an occurrence of εx (which does not begin a subterm which has just been numbered, or else you should have gone on to (3)!), assign j to that occurrence of x and to every other occurrence of x which is bound by that occurrence of $\forall x$, $\exists x$, or εx. Go to (3). If you reach the end of t without finding anything to number, quit.

(3) Reset j to $j + 1$. Return to (2).

Now we can define the *ground term* of an ε–term t to be the ε–term which results if in place of all the numbered variables and subterms we put the corresponding variable v_j. So each ground term is an ε-term of form $\varepsilon v_1.\varphi$.

The first two clauses of the following proposition are obvious. The third holds because our definition of substitution is such that if t is not free for y in $\varepsilon x.\varphi$ the renaming of bound variables in $\varepsilon x.\varphi$ ensures that wherever y occurs free in $\varepsilon x.\varphi$, t will occur free in $\varepsilon x.\varphi[x/t]$.

Proposition 1.

(1) *Each ε–term t has a unique ground term.*

(2) *Each ε–term t arises out of a ground term by renaming bound variables and simultaneous substitution of terms for free variables.*

(3) $\varepsilon x.\varphi$ *and* $\varepsilon x.\varphi[y/t]$ *have the same ground terms.* □

We can now give the following definitions.

Definition 2. Let G be the set of ground terms of $\mathcal{L}$. Let D be a set, and let S be the set of finite sequences of members of D. Then a function $E : \mathcal{P}(D) \times G \times S \rightarrow D$ is a *Hilbertian choice function for* $\mathcal{L}$ *and* D if

$$E(M, t, r) \in M$$

when $M \neq \emptyset \subseteq D$, $t \in G$, and $r = \langle d_1, \ldots, d_n \rangle \in S$, where $n = card(\mathbf{FV}(t))$.

Definition 3. $M = \langle D, \mathcal{R}, \mathbf{c}, E \rangle$ is a classical Hilbertian ε–structure for $\mathcal{L}$ if $\langle D, \mathcal{R}, \mathbf{c} \rangle$ is a structure for $\mathcal{L}' \subseteq \mathcal{L}$, the ε–free portion of $\mathcal{L}$ (cf. Chapter 6), and E is a Hilbertian choice function for $\mathcal{L}$ and D.

We can leave the definitions of $\varrho, \varrho(x/d) : \mathbf{Var} \rightarrow D$ as before. However, we will obviously need to modify our definition of the interpretation of the ε-terms. For each ε-term t, we denote the ground term of t by '$\gamma(t)$'. To state the appropriate clause of our definition, we make use of the numbering described in Definition 1. This gives us a numbering of the free *occurrences* in t of the proper subterms of t which are members of $\mathbf{FT}(t)$. Let $\langle t_1, \ldots, t_n \rangle$ be the sequence we get by listing these occurrences in the order of this numbering. There will then be n free variables in $\gamma(t)$. Let $\vec{x} = \langle x_1, \ldots, x_n \rangle$ be the sequence of free variables of $\gamma(t)$. If $\gamma(t) \equiv \varepsilon v_1.\beta$, put

$$M' = \{\, d \in D \mid \|\beta\|_{\mathcal{M}}^{\varrho(v_1/d)(x_1/[t_1]_{\mathcal{M}}^{\varrho})\ldots(x_n/[t_n]_{\mathcal{M}}^{\varrho})} = 1 \,\}.$$

We now have the tools we need to extend the definitions of $[t]_{\mathcal{M}}^{\varrho} \in D$ and $\|\varphi\|_{\mathcal{M}}^{\varrho} \in$ **2**.

Definition 4. To extend the definitions of $[t]^{\varrho}_{\mathcal{M}} \in D$ and $\|\varphi\|^{\varrho}_{\mathcal{M}} \in \mathbf{2}$ from the terms and formulas of $\mathcal{L}'$ to all of **Terms** and **Wffs**, we can rewrite Definition 6.1, except that we write $\mathcal{L}$ for $\mathcal{L}'$, and we add the following clause to it, using the notation introduced in the preceding paragraph:

$$[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = E\,(M', \gamma(\varepsilon x.\varphi), \langle [t_1]^{\varrho}_{\mathcal{M}}, \ldots, [t_n]^{\varrho}_{\mathcal{M}} \rangle)\,.$$

The next step is to make sure that our definition of substitution fits together appropriately with our semantic definitions.

Theorem 1. *Let $\mathcal{M}$ be a classical Hilbertian ε-structure for $\mathcal{L}$, M a well-formed expression of $\mathcal{L}$, and ϱ, ϱ' two valuations in $\mathcal{M}$ such that for any $x \in \mathbf{FV}(M)$, $\varrho(x) = \varrho'(x)$. Then*

(1) $M \equiv t \in \mathbf{Terms} \implies [t]^{\varrho}_{\mathcal{M}} = [t]^{\varrho'}_{\mathcal{M}}\,.$

(2) $M \equiv \varphi \in \mathbf{Wffs} \implies \|\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho'}_{\mathcal{M}}\,.$

Proof. We can use the proof of Theorem 7.1. We only need to note in the last step that the inductive hypothesis gives us that $\varrho(x/d)$ and $\varrho'(x/d)$ agree on the interpretations of $\mathbf{FT}(\varphi)$. $\square$

Theorem 2. *Let M be a well-formed expression of $\mathcal{L}$, $\mathcal{M}$ a classical Hilbertian ε-structure for $\mathcal{L}$, ϱ a valuation in $\mathcal{M}$, and $s \in$ **Terms**. Then*

(1) $M \equiv t \in \mathbf{Terms} \implies [t]^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})}_{\mathcal{M}} = [t[x/s]]^{\varrho}_{\mathcal{M}}.$

(2) $M \equiv \varphi \in \mathbf{Wffs} \implies \|\varphi\|^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})}_{\mathcal{M}} = \|\varphi[x/s]\|^{\varrho}_{\mathcal{M}}.$

Proof. Once again we can use essentially the same proof we used to prove the corresponding theorem for the classical extensional case (i. e. Theorem 7.2). We only need to note in the parts of the proof that refer to the interpretation of ε-terms that the inductive hypotheses guarantee that the free occurrences of proper subterms get the same interpretations. □

A digression. Note that the analogue of Lemma 7.3 holds only because we have been careful in our definition of *subformula* and in our statement of that lemma. We have stated clause (3) as follows:

(3) If β is a subformula of φ, and φ' results from φ by replacing 0 or more occurrences of β by β', then for any $\varrho, \mathcal{M}$, $\|\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi'\|^{\varrho}_{\mathcal{M}}$,

where we are assuming that $\|\beta\|^{\varrho}_{\mathcal{M}} = \|\beta'\|^{\varrho}_{\mathcal{M}}$ for all ϱ, $\mathcal{M}$.

Now, for the extensional calculus we can actually prove something stronger, namely:

Proposition 2. *If β is a subformula of φ or $\varepsilon x.\beta$ occurs in φ, and φ' results from φ by replacing 0 or more occurrences of β by β' or $\varepsilon x.\beta$ by $\varepsilon x.\beta'$, then if, for any ϱ, $\mathcal{M}$, $\|\beta\|^{\varrho}_{\mathcal{M}} = \|\beta'\|^{\varrho}_{\mathcal{M}}$, then*

$$\|\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi'\|^{\varrho}_{\mathcal{M}}$$

for any $\mathcal{M}$ and ϱ also.

Proof. It is obviously enough to show

$$\|P_i t_1 \dots t_n\|^{\varrho}_{\mathcal{M}} = \|P_i t_1 \dots t_{n-1} t'_n\|^{\varrho}_{\mathcal{M}},$$

where $t_n = \varepsilon x.\beta$ and $t'_n = \varepsilon x.\beta'$. But $\|\beta\|^{\varrho(x/d)}_{\mathcal{M}} = \|\beta'\|^{\varrho(x/d)}_{\mathcal{M}}$ for any $d \in D$, so $\|\forall x.(\beta \Leftrightarrow \beta')\|^{\varrho}_{\mathcal{M}} = 1$ for any ϱ, $\mathcal{M}$, and so $[\varepsilon x.\beta]^{\varrho}_{\mathcal{M}} = [\varepsilon x.\beta']^{\varrho}_{\mathcal{M}}$, by Lemma 7.10.

So

$$\mathcal{R}_i(\langle [t_1]^{\varrho}_{\mathcal{M}}, \ldots, [t_n]^{\varrho}_{\mathcal{M}} \rangle) = \mathcal{R}_i(\langle [t'_1]^{\varrho}_{\mathcal{M}}, \ldots, [t'_n]^{\varrho}_{\mathcal{M}} \rangle) . \quad \Box$$

Obviously this proof depends crucially on Lemma 7.10, which in turn depends on (Ack). From Lemma 7.3 and Proposition 2, the Soundness and Completeness Theorems, and the obvious corollaries of Lemma 7.11 that for all ϱ, $\mathcal{M}$

$$\|\forall x.\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi[x/\varepsilon x.\neg\varphi]\|^{\varrho}_{\mathcal{M}} \qquad \text{and} \qquad \|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi[x/\varepsilon x.\varphi]\|^{\varrho}_{\mathcal{M}} ,$$

it follows that the universal and existential quantifiers are eliminable from the classical extensional ε-calculus. So we can formulate that calculus equivalently without taking the quantifiers as basic.

On the other hand, in the case of the classical Hilbertian ε-calculus, while we can still prove the logical equivalences just mentioned, we can no longer prove Proposition 2. The problem is that while, for example, $\forall y.\varphi$ and $\varphi[y/\varepsilon y.\neg\varphi]$ are equivalent formulas, $\varepsilon x.(\forall y.\varphi)$ has a different ground term from $\varepsilon x.(\varphi[y/\varepsilon y.\neg\varphi])$, and so possibly

$$[\varepsilon x.(\forall y.\varphi)]^{\varrho}_{\mathcal{M}} \neq [\varepsilon x.(\varphi[y/\varepsilon y.\neg\varphi])]^{\varrho}_{\mathcal{M}} .$$

So we would have a slightly different calculus if we had *defined* the quantifiers, rather than taking them as primitives. (Asser seems to have been the first to point this out, in [Asser, p. 62].)

To return to the main thread of this section, we can prove the basic metatheorems for the Hilbertian calculus by slightly modifying the proofs we used in the extensional case. First we define

Definition 5. A *Hilbertian classical ε-Hintikka set* in $\mathcal{L}$ is a set Ψ of formulas which meets conditions (1)-(16) of the definition of a classical extensional ε-Hintikka set (Definition 7.2) and such that

(17) Let $t_1, \ldots, t_n, s_1, \ldots, s_n \in$ **Terms**, $\vec{t} = \langle t_1, \ldots, t_n \rangle$, and $\vec{s} = \langle s_1, \ldots, s_n \rangle$. Then if for $i = 1, \ldots, n$, $t_i = s_i \in \Psi$,

$$(\varepsilon y.\varphi)[\vec{x}/\vec{t}] = (\varepsilon y.\varphi)[\vec{x}/\vec{s}] \in \Psi.$$

Obviously, we now want to show that every such set is satisfiable. Of course, the main difficulty involved in doing so is finding an appropriate Hilbertian choice function.

We begin by defining $[t]$ and $\mathcal{R}_i$ just as we did in Chapter 7. We also use Definition 7.5 exactly as stated. However, since we have changed clause (17) of the definition of Ψ, we cannot simply define E as we did in Chapter 7. In Chapter 7 we showed that if $\langle \varphi, x, \vec{s} \rangle$ and $\langle \beta, y, \vec{t} \rangle$ both represent N, then $(\varepsilon x.\varphi)[\vec{z}/\vec{s}] \sim (\varepsilon y.\beta)[\vec{z}/\vec{t}]$. We can't prove this in the absence of (Ack). What we can show is

Lemma 1. *Let $\langle s_1, \ldots, s_n \rangle$ be the sequence of occurrences in $\varepsilon x.\varphi$ of proper subterms of $\varepsilon x.\varphi$ which are also members of* **FT**$(\varepsilon x.\varphi)$*. Let $\langle t_1, \ldots, t_n \rangle$ be the corresponding set for $\varepsilon y.\beta$. Suppose that $s_i = t_i \in \Psi$ for $i = 1, \ldots, n$, and that $\gamma(\varepsilon y.\beta) = \gamma(\varepsilon x.\varphi) = \varepsilon v_1.\alpha$. Then $\varepsilon y.\beta = \varepsilon x.\varphi \in \Psi$.*

Proof. By (17), $(\varepsilon v_1.\alpha)[\vec{x}/\vec{s}] = (\varepsilon v_1.\alpha)[\vec{x}/\vec{t}]$, where $\vec{x}$ is the sequence of free variables in $\varepsilon v_1.\alpha$. But $(\varepsilon v_1.\alpha)[\vec{x}/\vec{s}]$ is an alphabetic variant of $\varepsilon x.\varphi$, and $(\varepsilon v_1.\alpha)[\vec{x}/\vec{t}]$ is an alphabetic variant of $\varepsilon y.\beta$. It is not hard to see that if s is an alphabetic variant of s' and $s = t \in \Psi$, then $s' = t \in \Psi$. So by two applications of this fact, $\varepsilon x.\varphi = \varepsilon y.\beta \in \Psi$. □

Definition 6. We can define our Hilbertian choice function $E : \mathcal{P}(D) \times G \times S \to D$ by stipulating that if $M \subseteq D$ is represented by $\langle \varphi, y, \vec{t} \rangle$ and, writing $\vec{s} = \langle s_1, \ldots, s_n \rangle$ for the sequence of free occurrences in $(\varepsilon y.\varphi)[\vec{z}/\vec{t}]$ of its proper subterms which are also members of $\mathbf{FT}((\varepsilon y.\varphi)[\vec{z}/\vec{t}])$, we can put

$$E(M, \gamma((\varepsilon y.\varphi)[\vec{z}/\vec{t}]), \langle [s_1], \ldots, [s_n] \rangle) = \left[(\varepsilon y.\varphi)[\vec{z}/\vec{t}] \right].$$

If M is not representable, or if $\varepsilon v_1.\varphi \in G$ is such that $\langle \varphi, v_1, s \rangle$ does not represent M, or if s is a sequence which is too long or too short, then we can let $E(M, \varepsilon v_1.\varphi, s)$ be an arbitrary element of D. (Note that we have defined Hilbertian choice functions in such a way that they need not be choice functions at all on the "don't cares".)

We can take as our Hilbertian ε-structure $\mathcal{M} = \langle D, \{ \mathcal{R}_i \mid i \in I \}, \mathbf{c}, E \rangle$, where D is the set of $\sim$-equivalence classes of **Terms**, the $\mathcal{R}_i$ and $\mathbf{c}$ are defined as in Chapter 7, and E is the Hilbertian choice function just defined.

Once again we will write $\mathbf{t}$ for a representative of $[t]$, and we will write, for $\vec{x} = \langle x_1, \ldots, x_n \rangle$ and ϱ a valuation in $\mathcal{M}$, $\rho(\mathbf{x})$ for $\langle \rho(\mathbf{x_1}), \ldots, \rho(\mathbf{x_n}) \rangle$.

Lemma 2. *Let $\mathcal{M}$ be a well-formed expression of $\mathcal{L}$ such that* $\mathbf{FV}(M) \subseteq \{x_1, \ldots, x_n\}$. *Then for any valuation ϱ in $\mathcal{M}$,*

(1) $M \equiv t \in \mathbf{Terms} \implies [t]^{\varrho}_{\mathcal{M}} = [t[\vec{x}/\rho(\mathbf{x})]]$

(2) $M \equiv \varphi \in \mathbf{Wffs} \implies (\varphi[\vec{x}/\rho(\mathbf{x})] \in \Psi \implies \|\varphi\|^{\varrho}_{\mathcal{M}} = 1)$

(3) $M \equiv \varphi \in \mathbf{Wffs} \implies (\neg\varphi[\vec{x}/\rho(\mathbf{x})] \in \Psi \implies \|\varphi\|^{\varrho}_{\mathcal{M}} = 0)$

Proof. We can use the proof of Lemma 7.7, provided we change the clause for ε-terms. So suppose $N = \{ [t] \in D \mid \|\beta\|_{\mathcal{M}}^{\varrho(y/[t]^{\varrho}_{\mathcal{M}})} = 1 \}$. By IH, $N = \{[t] \in D \mid \beta[\vec{x}/\boldsymbol{\rho}(\mathbf{x}), y/t] \in \Psi \}$. So $\langle \beta, y, \boldsymbol{\rho}(\mathbf{x})\rangle$ represents N. If we write $\langle s_1, \ldots, s_l\rangle$ for the series of free occurrences of proper subterms of $(\varepsilon y.\beta)[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]$,

$$[\varepsilon y.\beta]^{\varrho}_{\mathcal{M}} = E\left(N, \gamma\left((\varepsilon y.\beta)[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]\right), \langle [s_1], \ldots, [s_l]\rangle\right)$$

$$= (\varepsilon y.\beta)[\vec{x}/\boldsymbol{\rho}(\mathbf{x})]. \quad \square$$

We can conclude that the following theorem holds.

Theorem 3. *Any (classical Hilbertian) ε-Hintikka set is satisfiable.* □

The point of the long song and dance used to prove Theorem 7.3 was that it allowed us to prove Soundness, Completeness and Compactness for the classical extensional ε-calculus. Theorem 3 will allow us to do the same for the Hilbertian case in much the same manner.

We can define in the obvious way the appropriate notion of *finitary semantic closure* for this semantics. If we write '$\Gamma \vdash_{\varepsilon} \varphi$' for "$\varphi$ is derivable from Γ in the classical Hilbertian ε-calculus", we can also define *deductive closure* for this calculus.

The analogues of Lemmas 7.9(a) and 7.9(b) can be proved for these closure operations. However, as we have already mentioned, Lemmas 10(a) and (b) will obviously no longer be available to us. Of course, we won't need them in our later proofs, either. Instead, we will need

Lemma 3. (1) $[t_1]^{\varrho}_{\mathcal{M}} = [t_2]^{\varrho}_{\mathcal{M}} \implies [\varepsilon x.\varphi[y/t_1]]^{\varrho}_{\mathcal{M}} = [\varepsilon x.\varphi[y/t_2]]^{\varrho}_{\mathcal{M}}$.

(2) $\Gamma \vdash_\varepsilon t_1 = t_2 \implies \Gamma \vdash_\varepsilon \varepsilon x.\varphi[y/t_1] = \varepsilon x.\varphi[y/t_2]$.

Proof. (2) is obvious by the $(A6)$ axiom For (1), suppose $[t_1]^\varrho_{\mathcal{M}} = [t_2]^\varrho_{\mathcal{M}}$. By Theorem 2(1) twice,

$$[\varepsilon x.\varphi[y/t_1]]^\varrho_{\mathcal{M}} = [\varepsilon x.\varphi]^{\varrho(y/[t_1]^\varrho_{\mathcal{M}})}_{\mathcal{M}} = [\varepsilon x \varphi]^{\varrho(y/[t_2]^\varrho_{\mathcal{M}})}_{\mathcal{M}} = [\varepsilon x.\varphi[y/t_2]]^\varrho_{\mathcal{M}}. \quad \square$$

We can prove, exactly as before, Lemmas 7.11(a) and (b). And so we have

Theorem 4 (Soundness Theorem). *Let* $\Gamma \subseteq$ **Wffs**, $\varphi \in$ **Wffs**. *Then*

$$\Gamma \vdash_\varepsilon \varphi \implies \Gamma \models \varphi.$$

Proof. By Theorem 6.1, we need only show that (ε), $(A6)$, and (α) are valid schemes in Hilbertian ε–structures. For (ε) and (α) the proof is the same as in the proof of Theorem 7.5, and $(A6)$ follows from Lemma 3((1)). $\square$

We can use the proofs of Theorems 7.6 and 7.7 to prove the following two theorems.

Theorem 5. (1) *If* $\mathbf{f} \notin \mathbf{C}_s(X)$, *then there is a* Y *such that* $X \subseteq Y$ *and* $Max_{\mathbf{C}_s}(Y)$.

(2) *If* $\mathbf{f} \notin \mathbf{C}_d(X)$, *then there is a* Y *such that* $X \subseteq Y$ *and* $Max_{\mathbf{C}_d}(Y)$. $\square$

Theorem 6. (1) $Max_{\mathbf{C}_s}(X) \implies (\varphi \in X \iff \neg\varphi \notin X)$.

(2) $Max_{\mathbf{C}_d}(X) \implies (\varphi \in X \iff \neg\varphi \notin X)$. $\square$

However, a slight modification is needed in the next case.

Theorem 7. *For* $i \in \{d, s\}$,

$Max\,\mathbf{C}_i(X) \implies X$ *is a classical Hilbertian* ε*–Hintikka set.*

Proof. The proof is just the same as the proof of Theorem 7.8, except for clauses (16) and (17). For (16), with $i = s$, we also need to point out that $\gamma(\varepsilon x.\varphi) \equiv \gamma(\varepsilon x.\varphi[x/y])$, and that the interpretations of the appropriate free terms in the two terms will be the same for any ϱ and $\mathcal{M}$.

For our new version of (17), the case of $i = d$ follows from the $(A6)$ schema, the fact that simultaneous substitution is defined in terms of sequential substitution, and the obvious fact that if $\Gamma \vdash_\varepsilon \varphi$, then $\Gamma \vdash_\varepsilon \varphi'$ if φ' is an alphabetic variant of φ. For $i = s$, the result follows by Lemma 3(1), the fact that alphabetic variants are semantically equivalent, and the fact that simultaneous substitution is defined in terms of sequential substitution. □

Thus we have

Theorem 8 (Compactness Theorem). *If* $X \subseteq$ **Wffs** *is such that every finite subset of* X *is satisfiable, then* X *is satisfiable in some* $\mathcal{M}$ *such that* $card(\mathcal{M}) \leqslant card(\mathcal{L})$. □

Theorem 9 (Completeness Theorem).

(1) *If* Γ *is consistent, then* Γ *is satisfiable in some model such that* $card(\mathcal{M}) \leqslant card(\mathcal{L})$.

(2) $\Gamma \models \varphi \implies \Gamma \vdash_\varepsilon \varphi$. □

We could obviously go on from here to prove that the ε-theorems, the Satisfiability Theorem, and so on, hold for the classical Hilbertian calculus.

Remark. The (Ack) principle is an extensionality condition in a rather obvious sense. However, (A6) is also a species of extensionality condition. Indeed, we might describe the effect of the move from the extensional calculus to the Hilbertian calculus as one of pushing the requirement that the denotation of an ε–term be determined extensionally down one level in the following sense. Clearly it is no longer the case that the fact that two predicates are coextensive guarantees that they determine the same *ideal* or *paradigm* object. So, to use Bell's example once again, it is possible that the ideal human being is different from the ideal featherless biped. However, *embedded* ε–terms are treated extensionally insofar as they determine the denotation of the ε–terms in which they are embedded. So if we take a two place predicate interpreted as "x is the best thing to do with y", and it so happens that my pet cat Piggy is both the ideal pet and the ideal candidate for roasting for supper, then (A6) seems to ensure that either the ideal thing to do with the ideal pet is to eat her for supper, or that the ideal thing to do with the ideal candidate for roasting for supper is to scratch her ears.

A natural question is whether we can give a semantics for the ε–calculus which does not make any such extensionality assumption. We turn to this task now.

2. Semantics for the Classical ε–Calculus

Once again we will describe the new semantics in terms of the modifications it requires in the semantics already described. However, the treatment here will be even more cursory.

Definition 7 (Skeleton terms). In place of the ground terms of the last section, here we will assign to each ε–term of $\mathcal{L}$ a *skeleton term*. Again, we will begin by assigning numbers to some of the parts of the term, this time just to the variables.

(1) Set $j = 1$.

(2) Starting at the left, find the first unnumbered occurrence of a variable. If you reach the end of t without finding any, quit. If the first unnumbered occurrence of a variable occurs in a substring of form $\forall x$, $\exists x$, or εx, assign j to that occurrence of x and to all occurrences of x bound by that occurrence of $\forall x$, $\exists x$, or εx. If the first unnumbered occurrence of a variable is a free occurrence, assign j to that occurrence.

(3) Reset $j = j + 1$, and return to (2).

We can now define the *skeleton term* for an ε–term t to be the ε–term which results when each occurrence of any variable x in t is replaced by the variable v_j, where j is the number assigned to that occurrence of x by the numbering just described.

Definition 8. An *ε–choice function for $\mathcal{L}$ and D* is a map $E : \mathcal{P}(D) \times K \times S \to D$, where $D \neq \emptyset$ is a set, K is the set of skeletons of ε–terms of $\mathcal{L}$, and S is the set of finite sequences of elements of D, and E is such that

$$E(M, \varepsilon v_1.\beta, s) \in M$$

if $M \neq \emptyset \subseteq D$, $\varepsilon v_1.\beta \in K$, and $s = \langle d_1, \ldots, d_r \rangle$ where $r = card(\mathbf{FV}(\varepsilon v_1.\beta))$.

Definition 9. A *classical ε–structure for $\mathcal{L}$* is a quadruple $\mathcal{M} = \langle D, \mathcal{R}, \mathbf{c}, E \rangle$, where $\langle D, \mathcal{R}, \mathbf{c}\rangle$ is a structure for $\mathcal{L}' \subseteq \mathcal{L}$, the ε–free sublanguage of $\mathcal{L}$ (cf. Chapter 6), and E is an ε–choice function.

The clause which defines the interpretation of ε–terms by the choice function is simpler in this case than in the Hilbertian case.

For any ε–term t, denote by '$\kappa(t)$' the skeleton of t.

Again, we can extract an ordering of the free *occurrences* of variables in $\varepsilon x.\varphi$ from the above numbering, and these occurrences correspond exactly to the free variables of $\kappa(\varepsilon x.\varphi)$. Write $\langle y_1, \ldots, y_n\rangle$ for the free occurrences of variables in $\varepsilon x.\varphi$, and $\langle x_1, \ldots, x_n\rangle$ for the sequence of free variables of $\kappa(\varepsilon x.\varphi)$. If $\kappa(\varepsilon x.\varphi) \equiv \varepsilon v_1.\beta$, let

$$M = \{\, d \in D \mid \|\beta\|_{\mathcal{M}}^{\varrho(v_1/d)(x_1/\varrho(y_1))\ldots(x_n/\varrho(y_n))} = 1 \,\}.$$

We can then proceed as follows.

Definition 10. To extend the definition of $[t]^{\varrho}_{\mathcal{M}} \in D$ and $\|\varphi\|^{\varrho}_{\mathcal{M}}$ from the terms and formulas of $\mathcal{L}'$ to all of **Terms** and **Wffs** we can rewrite definition 6.1, except that we write $\mathcal{L}$ for $\mathcal{L}'$, and we add the following clause to it, using the notation introduced in the preceding paragraph:

$$[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = E\,(M, \kappa(\varepsilon x.\varphi), \langle \varrho(y_1), \ldots, \varrho(y_n)\rangle)\,.$$

This definition makes it particularly obvious that our first bookkeeping theorem holds.

Theorem 10. *Let $\mathcal{M}$ be a classical ε–structure for $\mathcal{L}$, M a well-formed expression of $\mathcal{L}$, and ϱ, ϱ' two valuations in $\mathcal{M}$ such that for any $x \in$* **FV**(M)*, $\varrho(x) = \varrho'(x)$.*

Then

(1) $M \equiv t \in \textbf{Terms} \implies [t]^{\varrho}_{\mathcal{M}} = [t]^{\varrho'}_{\mathcal{M}}$

(2) $M \equiv \varphi \in \textbf{Wffs} \implies [\varphi]^{\varrho}_{\mathcal{M}} = [\varphi]^{\varrho'}_{\mathcal{M}}$. □

However, our second bookkeeping theorem no longer holds in full generality. For example, it is now possible that

$$[\varepsilon x.Pxy]^{\varrho(y/[\varepsilon z.Rz]^{\varrho}_{\mathcal{M}})}_{\mathcal{M}} \neq [\varepsilon x.Px\varepsilon z.Rz]^{\varrho}_{\mathcal{M}} = [\varepsilon x.Pxy[y/\varepsilon z.Rz]]^{\varrho}_{\mathcal{M}}$$

because $\varepsilon x.Pxy$ and $\varepsilon x.Px\varepsilon z.Rz$ have different skeletons.

We also note that this also has as a consequence that the axioms $(Q5)$ and $(Q6)$ are not valid as originally stated. This is, of course, why we have described the classical ε–calculus as containing instead

(Q5′) $\forall x.\alpha \Rightarrow \alpha[x/t]$ provided t is free for x in α, and provided x does not have a free occurrence in the scope of an ε–bound variable.

(Q6′) $\alpha[x/t] \Rightarrow \exists x.\alpha$ provided t is free for x in α, and provided x does not have a free occurrence in the scope of an ε–bound variable.

For if we do not include the extra qualification in these axioms we run afoul of the fact that we can easily construct interpretations such that

$$\|\forall x.(\varepsilon y.y = b) \neq (\varepsilon y.y = x)\|^{\varrho}_{\mathcal{M}} = 1$$

(for example, we can get this result by ensuring that E is such that $E((M, \varepsilon v_1.v_1 = b, s) \neq E(M, \varepsilon v_1.v_1 = v_2, s)$ for all M and s) and yet, obviously,

$$\|(\varepsilon y.y = b) \neq (\varepsilon y.y = b)\|^{\varrho}_{\mathcal{M}} \neq 1.$$

And conversely this will obviously also be an interpretation such that we cannot conclude from

$$\|(\varepsilon y.y = b) = (\varepsilon y.y = b)\|_{\mathcal{M}}^{\varrho} = 1$$

that

$$\|\exists x.[(\varepsilon y.y = x) = (\varepsilon y.y = b)\|]_{\mathcal{M}}^{\varrho} = 1.$$

(Note also that substitutivity of identicals does not hold in general, which is clear from the fact that (A6) is not valid.)[1] It is easy to see that we can get around our present problem by making a simple modification to our bookkeeping theorem.

Theorem 11. *Let M be a well-formed expression of $\mathcal{L}$, $\mathcal{M}$ a structure for $\mathcal{L}$, ϱ a valuation in $\mathcal{M}$, and $s \in$* **Terms**. *Then*

(1) $M \equiv t \in$ **Terms** $\Longrightarrow [t]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = [t[x/s]]_{\mathcal{M}}^{\varrho}$

(2) $M \equiv \varphi \in$ **Wffs** $\Longrightarrow \|\varphi\|_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = \|\varphi[x/s]\|_{\mathcal{M}}^{\varrho}$

provided that x does not have a free occurrence in M which is a substring of an ε–term which is a proper subterm of M. □

In the part of the proof that corresponds to Lemma 7.3, the bolded proviso allows us to circumvent the failure warned about in the proof of that lemma. If we did not have the restriction in place, it would be possible that the substitution would change the skeleton. With the restriction in place, this cannot happen, so the result continues to hold. Of course, it holds precisely because it is a much weaker result with the proviso in place. On the other hand, it is still strong enough to allow us

[1] [Corcoran and Herring] raise these points as objections to the semantics for term forming operators presented in [Hatcher 1968].

to prove the result corresponding to Lemma 7.4, and so also Theorem 11, which is a weakened version of Theorem 7.2.

Definition 11. A *classical ε–Hintikka set* is a subset of **Wffs** satisfying clauses (1)–(16) of Definition 7.2.

Now in defining the canonical interpretation for an ε–Hintikka set, it is useful to leave the definition of representable sets unchanged, and to specify: If $N \subseteq D$ is represented by $\langle \varphi, y, \vec{t} \rangle$, and if $\langle x_1, \ldots, x_n \rangle$ is the sequence of free occurrences of variables in $(\varepsilon y.\varphi)[\vec{z}/\vec{t}]$, then we can define E by

$$E\left(N, \kappa\left((\varepsilon y.\varphi)[\vec{z}/\vec{t}]\right), \langle [x_1], \ldots, [x_n] \rangle\right) = \left[(\varepsilon y.\varphi)[\vec{z}/\vec{t}]\right].$$

Clearly $\left[(\varepsilon y.\varphi)[\vec{z}/\vec{t}]\right] \in N$, by (13). If N is not representable, or if $t \equiv \varepsilon y.\varphi$ with free variables $\vec{x} = \langle x_1, \ldots, x_n \rangle$ and $\langle \varphi, y, \vec{x} \rangle$ does not represent N, or if s is a sequence other than $\langle [x_1], \ldots, [x_n] \rangle$ then we can let $E(N, k, s)$ be an arbitrary $d \in D$.

We now have an ε–choice function E for $\mathcal{L}$ and D, and so an ε–structure for $\mathcal{L}$. The analogue of Lemma 7.7 is even easier than it was in the Hilbertian case. We only need to check the clause for ε–terms; but now $[\varepsilon y.\beta]^{\varrho}_{\mathcal{M}} = [(\varepsilon y.\beta)[x/\boldsymbol{\rho}(\mathbf{y})]]$ by definition of E. So we can get

Theorem 12. *Any (classical) ε–Hintikka set is satisfiable.* □

The rest of our metatheoretical investigations can continue very much as in the earlier cases. Obviously, we cannot prove anything like Lemmas 7.10(a) and 7.10(b), nor even Lemmas 3(a) and 3(b), but this is no problem since the clauses where these lemmas get used in the proof of the Soundness Theorems and in Theorems 8 and 7.8 no longer need to be proved.

Finally, we should note that the proof of the analogue of Lemma 7.11(a) depends only on the part of Theorem 7.2(2), which still holds in full generality.

So the proofs for the Soundness, Compactness, and Completeness Theorems, and so for the other metatheoretical theorems, including the two ε–theorems, go through much as before. Indeed, they are actually easier, since for the most part we need only drop clauses from the earlier proofs.

CHAPTER IX

BOOLEAN VALUED SEMANTICS FOR THE ε—CALCULUS

We began our discussion of the semantics of the ε-calculus with the classical, extensional case in Chapter 7. In the last chapter we looked at what was required to eliminate the assumption of extensionality, in part because we will need to avoid this assumption in our discussion of the intuitionistic case below. We will now generalize our treatment of the semantics for the classical ε-calculus in a different way, namely by allowing our lattice of truth-values to be something other than **2**. Once again, this is something that will need to be done below, anyway, for us to be able to deal with the intuitionistic case. We will also need to make some adjustments to the (Ack)-schema and the (*)-rule for us to be able to prove completeness in the Boolean valued case, and the same modifications will be needed in the intuitionistic case. These changes are required, as we shall see, by our desire to continue using, in as straightforward a way as possible, a choice function to interpret the ε-terms in combination with allowing more than two truth-values.

Before we get around to Boolean valued semantics for ε, a brief sketch of Boolean valued semantics for the first order predicate calculus with identity will be useful. We will also give a quick sketch of a Lindenbaum algebra proof of completeness for this semantics. This will help us keep straight which changes to the semantics of Chapter 6 are required by the move to Boolean valued semantics and which are due to ε, as well as giving us something to refer to in giving completeness proofs in §2 of this chapter and in later chapters.

1. Boolean Valued Structures for the First Order Predicate Calculus

The basic modification which needs to be made in moving from the structures for the first order predicate calculus described in Chapter 6 to Boolean valued structures is obviously the replacement of **2** as the lattice of truth values by an arbitrary Boolean algebra B (which we shall insist must be a *complete* Boolean algebra to ensure without fuss that our interpretations of the quantifiers are well-defined). However, it is also necessary to specify a map $\mathrm{eq}_{\mathcal{M}} : D \times D \to B$ to interpret the identity symbol. For if we simply leave the interpretation as before, we are allowing formulas to take as truth-values any element of B, except formulas of form $t_1 = t_2$, which must take as values either 0 or 1. We will also need to specify some further conditions that the map $\mathrm{eq}_{\mathcal{M}}$ must meet so that identity remains well behaved—that is, to ensure that axioms I1-I3 remain valid.

So let $\mathcal{L}'$ be, as usual, the ε–free sublanguage of $\mathcal{L}$. We assume that the primitive predicates and constants of $\mathcal{L}'$ are given in the form of indexed sets $\{ P_i : i \in I \}$ and $\{ c_j : j \in J \}$, and for each $i \in I$ the arity of P_i is denoted by $\zeta(i)$. A *Boolean valued $\mathcal{L}'$-pre-structure* is a quintuple $\mathcal{M} = \langle D, B, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c} \rangle$ where

(1) $D \neq \emptyset$ is a set.

(2) B is a complete Boolean algebra.

(3) $\mathcal{R} : I \to \{ f \mid$ for some positive integer n, $f : D^n \to B \}$ is such that for each $i \in I$, $\mathcal{R}_i : D^{\varsigma(i)} \to B$.

(4) $\mathrm{eq}_{\mathcal{M}} : D \times D \to B$.

(5) $\mathbf{c} : J \to D$.

Definition 1. We can define the *interpretations,* $[t]^{\varrho}_{\mathcal{M}} \in D$ *and* $\|\varphi\|^{\varrho}_{\mathcal{M}} \in B$, *of* t *and* φ *in* $\mathcal{M}$ *under* ϱ by rewriting Definition 6.1, except that we replace the clause for = by

$$\|t_1 = t_2\|^{\varrho}_{\mathcal{M}} = \mathrm{eq}_{\mathcal{M}}\left([t_1]^{\varrho}_{\mathcal{M}}, [t_2]^{\varrho}_{\mathcal{M}}\right)$$

(and by noting that the symbols $\Rightarrow$, $\vee$, $\wedge$, and * on the right now refer to operations in B).

A Boolean valued $\mathcal{L}'$–pre-structure $\mathcal{M}$ is a *Boolean valued* $\mathcal{L}'$*-structure* if for all $t_1, t_2, t_3, t_4, \ldots, t_{2n}$ in the set of terms of $\mathcal{L}'$, and for all ϱ in $\mathcal{M}$,

(1) $\|t_1 = t_1\|^{\varrho}_{\mathcal{M}} = 1$

(2) $\|(t_1 = t_{n+1} \wedge \cdots \wedge t_n = t_{2n} \wedge Pt_1 \ldots t_n) \Rightarrow Pt_{n+1} \ldots t_{2n}\|^{\varrho}_{\mathcal{M}} = 1$

(3) $\|(t_1 = t_2 \wedge t_3 = t_4 \wedge t_1 = t_3) \Rightarrow t_2 = t_4\|^{\varrho}_{\mathcal{M}} = 1$.

If $\mathcal{M} = \langle D, B, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c}\rangle$ is a Boolean valued structure for $\mathcal{L}'$, we will also say that $\mathcal{M}$ is a *B–valued structure for* $\mathcal{L}'$. Note that the structures for $\mathcal{L}'$ as defined earlier are precisely the **2**–valued structures for $\mathcal{L}'$. If $\mathcal{M}$ is a B–valued structure for $\mathcal{L}'$ and there is some valuation in $\mathcal{M}$ such that $\|\varphi\|^{\varrho}_{\mathcal{M}} = 1$ we say that φ is *satisfiable* and that it is *satisfied (by* ϱ*)* in $\mathcal{M}$. If $\|\varphi\|^{\varrho}_{\mathcal{M}} = 1$ for all ϱ in $\mathcal{M}$, we say that φ is *valid in* $\mathcal{M}$. If φ is valid for all B–valued structures for $\mathcal{L}'$, we say that

φ is B-*valid*. We will temporarily write '$\Gamma \vdash \varphi$' for 'φ is derivable from Γ in the first order predicate calculus,' and for the remainder of this section we will abuse terminology and use **Wffs** and **Terms** to refer to the sets of formulas and terms of $\mathcal{L}'$ rather than of $\mathcal{L}$.

Suppose that Σ is a consistent set of *sentences* of $\mathcal{L}'$. We next describe how the Lindenbaum algebra for Σ is constructed. Define for any $\varphi, \psi \in \mathbf{Wffs}$,

$$\varphi \approx \psi \iff \Sigma \vdash \varphi \Leftrightarrow \psi.$$

This is obviously an equivalence relation. We will write $\|\varphi\|$ for $\{\, \psi \in \mathbf{Wffs} \mid \varphi \approx \psi \,\}$. We put $B = \{\, \|\varphi\| \mid \varphi \in \mathbf{Wffs} \,\}$, and order B by

$$\|\varphi\| \leqslant \|\psi\| \iff \Sigma \vdash \varphi \Rightarrow \psi.$$

This is evidently a sound definition. It is not hard to prove

Lemma 1. *$\langle B, \leqslant \rangle$ is a Boolean algebra. Moreover, $\|\varphi\| = 1 \iff \Sigma \vdash \varphi$ and $\|\varphi\| = 0 \iff \Sigma \vdash \neg\varphi$.*

Proof. The proof is standard. The relation $\leqslant$ is antisymmetric by definition, reflexive since $\varphi \Rightarrow \varphi$ is a propositional tautology, and transitive since it is a fact of propositional logic that if $\Sigma \vdash \varphi \Rightarrow \psi$ and $\Sigma \vdash \psi \Rightarrow \beta$, then $\Sigma \vdash \varphi \Rightarrow \beta$. So $\leqslant$ is a partial ordering on B.

By propositional logic, $\Sigma \vdash (\varphi \wedge \psi) \Rightarrow \varphi$, $\Sigma \vdash (\varphi \wedge \psi) \Rightarrow \psi$, and if $\Sigma \vdash \beta \Rightarrow \varphi$ and $\Sigma \vdash \beta \Rightarrow \psi$, then $\Sigma \vdash \beta \Rightarrow (\varphi \wedge \psi)$. It follows that $\|\varphi\| \wedge \|\psi\| = \|\varphi \wedge \psi\|$. The proof that $\|\varphi\| \vee \|\psi\| = \|\varphi \vee \psi\|$ is similar. Thus $\langle B, \leqslant \rangle$ is a lattice. Since

$$((\varphi \vee \psi) \wedge \beta) \Leftrightarrow ((\varphi \wedge \beta) \vee (\psi \wedge \beta))$$

is a propositional tautology, it is in fact a distributive lattice.

It is easy to see that $\|\beta \Rightarrow \beta\| = 1$ and $\|\neg(\beta \Rightarrow \beta)\| = 0$, and $0 \neq 1$ since Σ is consistent. If $\Sigma \vdash \varphi$, then for any $\psi \in$ **Wffs**, $\Sigma \vdash \psi \Rightarrow \varphi$, so $\|\psi\| \leqslant \|\varphi\|$ for all ψ. So $\|\varphi\| = 1$. Conversely, suppose $\|\varphi\| = 1$. Then $\|\beta \Rightarrow \beta\| \leqslant \|\varphi\|$, which is to say that $\Sigma \vdash (\beta \Rightarrow \beta) \Rightarrow \varphi$, so $\Sigma \vdash \varphi$ by modus ponens. So $\Sigma \vdash \varphi \iff \|\varphi\| = 1$. That $\Sigma \vdash \neg\varphi \iff \|\varphi\| = 0$ is shown similarly. Finally, since $(\varphi \vee \neg\varphi)$ and $\neg(\varphi \wedge \neg\varphi)$ are propositional tautologies for any $\varphi \in$ **Wffs**, the complement of $\|\varphi\|$ in $\langle B, \leqslant\rangle$ is $\|\neg\varphi\|$. $\square$

We can now give the standard proof of

Lemma 2. *For* $\varphi \in$ **Wffs**, $x \in$ **Var**, $t \in$ **Terms**,

$$\|\exists x.\varphi\| = \bigvee_{t \in \mathbf{Terms}} \|\varphi[x/t]\|$$

and

$$\|\forall x.\varphi\| = \bigwedge_{t \in \mathbf{Terms}} \|\varphi[x/t]\|\,.$$

Proof. We prove the second claim. It is well known that in the classical predicate calculus, for any $\varphi \in$ **Wffs**, $t \in$ **Terms**, $\vdash \forall x.\varphi \Rightarrow \varphi[x/t]$, so for any $t \in$ **Terms** $\Sigma \vdash \forall x.\varphi \Rightarrow \varphi[x/t]$. Thus $\|\forall x.\varphi\| \leqslant \|\varphi[x/t]\|$, i. e. $\|\forall x.\varphi\|$ is a lower bound for $\{\, \|\varphi[x/t]\| \mid t \in \mathbf{Terms} \,\}$.

Suppose $\|\beta\|$ is a lower bound for that set. Since at most finitely many variables occur in β and φ, we can find a z which occurs in neither. Since $\|\beta\|$ is a lower bound, $\|\beta\| \leqslant \|\varphi[x/z]\|$, i. e. $\Sigma \vdash \beta \Rightarrow \varphi[x/z]$. But then $\Sigma \vdash \beta \Rightarrow \forall z.\varphi[x/z]$ by

$\forall$-introduction, since z does not occur in β, and so $\Sigma \vdash \beta \Rightarrow \forall x.\varphi$ by renaming bound variables. So $\|\forall x.\varphi\|$ is an infimum, i. e. $\|\forall x.\varphi\| = \bigwedge_{t\in\mathbf{Terms}} \|\varphi[x/t]\|$.

The proof of the first claim is dual. ☐

B, so constructed, is the Lindenbaum algebra of Σ. We now have almost all the elements in place to provide a B–valued structure for $\mathcal{L}'$. We can, for $\mathcal{P}_i \in \mathbf{Pred}_n$, put

$$\mathcal{R}_i(t_1, \ldots, t_n) = \|P_i t_1 \ldots t_n\|,$$

for constants c_i put $\mathbf{c}(i) = c_i$, and for $\mathrm{eq}_{\mathcal{M}}$ we put

$$\mathrm{eq}_{\mathcal{M}}(t_1, t_2) = \|t_1 = t_2\|.$$

However, we have defined Boolean valued models as taking *complete* Boolean algebras as truth value lattices, and the Lindenbaum algebra need not be complete, so we need to appeal to

Proposition 1. *Every Boolean algebra B is isomorphic to a subalgebra of a complete Boolean algebra B', and the injection $i : B \to B'$ preserves all joins and meets.* ☐

For a proof, see [Rasiowa and Sikorski, pp. 88–91]. There are in general many such completions we could appeal to, so we will henceforth assume that when we appeal to Proposition 1 we have selected a particular i. We therefore can get a Boolean valued structure $\mathcal{M} = \langle \mathbf{Terms}, B', \{ i \circ \mathcal{R}_i : i \in I \}, i \circ \mathrm{eq}_{\mathcal{M}}, \mathbf{c}\rangle$, and Lemmas 1 and 2 will hold for B' also. Obviously we can take the identity map on

Var as a map π : **Var** $\to$ **Terms**, and so as a valuation in $\mathcal{M}$. We call this the canonical valuation of Σ.

It is straightforward to prove

Lemma 3. $\|\varphi\|^{\pi}_{\mathcal{M}} = 1 \iff \Sigma \vdash \varphi$. $\square$

This puts us in a position to prove

Theorem 1. *Let φ be a formula of $\mathcal{L}'$. The following are equivalent.*

(1) $\Sigma \vdash \varphi$

(2) $\|\varphi\|^{\pi}_{\mathcal{M}} = 1$ *in the canonical valuation of* Σ.

(3) $\Sigma \models \varphi$.

Proof sketch. For (1) $\implies$ (3), that is, to show soundness, it suffices to show two things. First, we must show that the axioms of the classical propositional and predicate calculi all are valid in Boolean valued models for arbitrary complete Boolean algebras. So consider (P1). The validity of (P1) follows from the fact that in any (complete) Boolean algebra B, for arbitrary $b, c \in B$, $b \Rightarrow (c \Rightarrow b) = 1$. The other axioms, both propositional and quantificational, are handled similarly. Secondly, we must show that (MP) is a valid rule of inference in Boolean-valued models, but this follows easily from the fact that $\alpha \wedge (\alpha \Rightarrow \beta) \leqslant \beta$ for all α and β in any Boolean algebra. Finally, the identity axioms are valid by our definition of Boolean valued structure.

(3) $\implies$ (2) is trivial, since $\mathcal{M}$ satisfies Σ. (2) $\implies$ (1) follows from Lemma 3. $\square$

The virtue of this theorem is that it gives us a proof of completeness and soundness which does not appeal directly to either the Axiom of Choice or the Ultrafilter Theorem. Of course, the price we pay is that we have to allow more than two truth values, even in classical logic. In addition, we need to detour through the theory of Boolean algebras long enough to prove Proposition 1 to ensure that our truth values belong to a complete Boolean algebra. (Alternatively, of course, we could complicate our definition of structure to allow Boolean algebras which are *sufficiently complete* in an obvious sense.) This leads some authors to skip Theorem 1. These authors go straight from the construction of the Lindenbaum algebra to factoring this algebra by an ultrafilter (as we are about to do) which, of course (cf. Proposition 2.4), gives us values in **2** for all formulas, and **2**, like all finite Boolean algebras, is complete.

So, suppose U is an ultrafilter in the Lindenbaum algebra $\langle B, \leqslant \rangle$. Then $\{ \varphi \mid \|\varphi\| \in U \}$ is a maximal consistent set (since U is an ultrafilter, either $\|\varphi\| \in U$ or $\|\neg\varphi\| \in U$ for every $\varphi \in \mathbf{Wffs}$, but not both or else $\|\varphi \wedge \neg\varphi\| = 0 \in U$). We define

$$t \sim t' \iff \|t = t'\| \in U.$$

This is easily shown to be an equivalence relation. We can now put $[t] = \{ t' \in \mathbf{Terms} \mid t \sim t' \}$ and define $\mathcal{T} = \{ [t] \mid t \in \mathbf{Terms} \}$. For $c_j \in \mathbf{Con}$, we put $\mathbf{c}(j) = [c_j]$. The following definition is easily seen to be independent of choice of representatives.

Let $P_i \in \mathbf{Pred}_n$. Then

$$\mathcal{R}_i([t_1], \ldots, [t_n]) = 1 \iff \|Pt_1 \ldots t_n\| \in U.$$

We now have all the elements in place for a $\mathbf{2}$–valued $\mathcal{L}'$–structure

$$\mathcal{M} = \langle \mathcal{T}, \mathbf{2}, \{ \mathcal{R}_i \mid i \in I \}, \{ c_j \mid j \in J \} \rangle.$$

(We suppress mention of $\mathrm{eq}_{\mathcal{M}}$ in this case, since it is simply the identity relation on $\mathcal{T}$.) Unfortunately, it is not in general true that $\mathcal{M}$ is a model of the set of sentences σ such that $\|\sigma\| \in U$. [Bell and Machover, p. 195] give the following counterexample.

Example 1. Let $\mathcal{L}'$ have no extralogical symbols. Let $D = \{a_0, a_1\}$, $\mathcal{M}$ be the unique structure with D as domain, and $\Sigma = \{\exists x.\exists y.x \neq y\}$. Let $\varrho : \mathbf{Var} \to D$ be the constant function such that $\varrho(x) = a_1$ for all $x \in \mathbf{Var}$. It is not hard to see that $U = \{ \|\varphi\| \in \langle B, \leqslant\rangle \mid \|\varphi\|^{\varrho}_{\mathcal{M}} = 1 \}$ is an ultrafilter in $\langle B, \leqslant\rangle$. But $\mathcal{M}' = \langle \mathcal{T}, \mathbf{2}, \emptyset, \emptyset\rangle$ is not a model of Σ since $\mathcal{T}$ has exactly one element.

However, this problem is not hard to circumvent. If we require that U meet the further condition that for any formula φ and any variable x

$$\|\exists x.\varphi\| \in U \iff \|\varphi[x/t]\| \in U$$

for some $t \in \mathbf{Terms}$, the set $\{\beta \mid \|\beta\| \in U\}$ is a *Henkin set*. We call an ultrafilter which meets this condition *perfect*. It is not hard to show (using methods similar to those used in proving Theorem 7.10)

Theorem 2. *Let U be a perfect ultrafilter in $\langle B, \leqslant\rangle$. Then if $t_1, \ldots, t_n \in$* **Terms**, *$\mathbf{FV}(\varphi) \subseteq \{x_1, \ldots, x_n\}$, and ϱ is a valuation in $\mathcal{M}$,*

$$\|\varphi\|^{\varrho(x_1/[t_1], \ldots, x_n/[t_n])}_{\mathcal{M}} = 1 \iff \|\varphi\,[x_1/t_1, \ldots, x_n/t_n]\| \in U.$$

It follows that $\mathcal{M}$ is an model of Σ. $\square$

The details of the proof can be found in [Bell and Machover, pp. 196–98].

Now, if we are willing to assume that $\mathcal{L}'$ is a countable language, then the family of joins $\{ \bigvee_{t \in \mathbf{Terms}} \|\varphi[x/t]\| \mid \varphi \in \mathbf{Wffs}, x \in \mathbf{Var} \}$ is countable. So by the Rasiowa–Sikorski Theorem, there is an ultrafilter U in $\langle B, \leqslant \rangle$ which respects that family of joins, which clearly is to say that U is perfect. We thus have another proof that every countable consistent set of sentences has a **2**–valued model, and so have a Completeness Theorem for such sets which depends on the Ultrafilter Theorem (via the Rasiowa–Sikorski Theorem), but not on the Axiom of Choice. However, this is a weaker theorem than our earlier completeness result, as another handy example lifted from [Bell and Machover] shows.

Example 2. Let $\mathcal{L}'$ have only *countably many* constants but *uncountably many* unary predicate symbols. If we let Σ be the set of sentences

$$\{ \exists x.P_i x \mid i \in I \} \cup \{ \forall x.\neg(P_i x \wedge P_j x) \mid i, j \in I, i \neq j \}$$

then any model of Σ must have an uncountable domain. Since $\mathcal{L}'$ has only countably many terms, this cannot happen in the structures for $\mathcal{L}'$ described above.

For the next theorem, then, if we restrict Σ to sets of sentences in countable languages, we have proved it by using the Rasiowa–Sikorski theorem. If we do not impose this restriction, we can appeal to our earlier completeness proof (which uses Zorn's Lemma) and the fact that **2**–valued structures are essentially just the structures described in chapter 6.

Theorem 3. *Let φ be a formula and Σ a set of sentences of $\mathcal{L}'$. Then the following claims are equivalent.*

(1) $\Sigma \vdash \varphi$.

(2) $\mathcal{M} \models \Sigma \implies \mathcal{M} \models \varphi$ *for all* **2**–*valued* $\mathcal{M}$.

(3) $\Sigma \models \varphi$. □

Corollary. *Let φ be a formula of $\mathcal{L}'$. Then the following claims are equivalent.*

(1) $\vdash \varphi$

(2) φ *is* **2**–*valid.*

(3) φ *is B–valid for some complete Boolean algebra B.*

(4) φ *is B–valid for all complete Boolean algebras B.* □

2. Boolean Valued Semantics for the ε–Calculus

2.1. Some (Hopefully) Intuitive Discussion.

As in the case of **2**–valued semantics, we can extend structures for the predicate calculus to structures for the ε–calculus by adding a choice function over the domain of interpretation to interpret the ε–terms—provided we are willing to adopt a strong extensionality principle. Indeed, we will need to assume a principle similar to, but slightly stronger than (Ack). However, this still leaves us with a modification which must be made to the formulation used in Definition 7.1. In the original formulation we put, following Asser and Leisenring,

$$[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = E\left(\{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 1 \,\}\right)$$

where E is a choice function which maps $\emptyset$ to an arbitrary fixed element of D. This is no longer appropriate. For it is easy to see that the (ε)–axiom scheme implies that we can prove

$$\exists x.\varphi \Leftrightarrow \varphi[x/\varepsilon x.\varphi]$$

in the ε–calculus. Once we open up the possibility that formulas can have, under a particular interpretation, truth-values other than 0 or 1, we open up the possibility that

$$\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} \neq 1,$$

hence

$$\{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 1 \,\} = \emptyset,$$

and yet

$$\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} \neq 0.$$

Under the original definitions, $[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}}$ is in this case an arbitrary but fixed element of D, and so possibly, e. g.,

$$[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} \in \{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 0 \,\}.$$

So the possibility exists that

$$\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} \neq \|\varphi[x/\varepsilon x.\varphi]\|^{\varrho}_{\mathcal{M}},$$

and so

$$\|\exists x.\varphi \iff \varphi[x/\varepsilon x.\varphi]\|^{\varrho}_{\mathcal{M}} \neq 1.$$

It is not hard to see how to get around this problem. If we put

$$[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = E\left(\{d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d\in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}}\}\right)$$

we ensure that $\exists x.\varphi \iff \varphi[x/\varepsilon x.\varphi]$ is valid, provided we also insist that for each φ and x the supremum $\bigvee_{d\in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}}$ is *attained*, i.e. is equal to $\|\varphi\|^{\varrho(x/d)}_{\mathcal{M}}$ for some $d \in D$.

It is perhaps worth noting that this modification arguably also gives us a closer approximation to the meaning of ε–terms suggested by the (ε) axiom itself. For that axiom suggests that $\varepsilon x.\varphi$ is an object which has the property φ if anything does. The Asser/Leisenring semantics for the extensional ε–calculus makes $\varepsilon x.\varphi$ an object which has φ if anything does, and an arbitrary object otherwise. The virtue claimed for this approach by both Leisenring and Asser is that it allows us to avoid the difficulties involved in allowing non-referring terms into our logical system. The approach works well enough provided we have only two truth values. But, as we have just seen, it causes problems in the Boolean valued case. The new definition, on the other hand, makes $\varepsilon x.\varphi$ an object which has the property φ *to the greatest extent possible*, or *which comes closest to making φ true*. Even in the **2**–valued case, this definition removes the impression of arbitrariness in the choice of $\varepsilon x.\varphi$ when φ is a property which no object happens to have, for the object assigned to $\varepsilon x.\varphi$ by the interpretation will remain one which has the property φ to the greatest extent possible.

On the other hand, if we are assuming (Ack) (or something similar), this definition will have as a consequence that universally instantiated properties and uninstantiated properties will have identical ε-terms. So we will need to add something like $(\varepsilon x.x = x) = (\varepsilon x.x \neq x)$ to our list of axioms. And it seems strange to say that the object self-identical to the greatest extent possible is also the object non-self-identical to the greatest extent possible.

However, [Hermes] gives a semantics for the ε-calculus (in the **2**-valued case) designed precisely to ensure that this condition is met (by stipulating that the choice function E must be such that $E(D) = E(\emptyset)$, rather than allowing E to assign an arbitrary element to the empty set). He does so in order to make his interpretation capture the spirit of some remarks made by Hilbert and Bernays in their introduction of ε in [Hilbert and Bernays, vol. 2]. In that work they initially define ε in terms of η, an operator which is designed to correspond to the indefinite article. The η symbol works like this: if $\exists x.\varphi$ is a theorem, then there is a term $\eta x.\varphi$, and we can infer $\varphi[x/\eta x.\varphi]$. However, it is rather inconvenient and arguably unnatural to have to deal with terms which can only be introduced provided some existential formula can be proved. So ε is introduced by defining

$$\varepsilon x.\varphi \equiv \eta x.(\exists y.\varphi[x/y] \Rightarrow \varphi).$$

Then $\varepsilon x.\varphi$ is a term for every x and φ, and every instance of the (ε) scheme is a theorem of the η-calculus. More to the present point, if $\varphi[x/t]$ is false for all $t \in$ **Terms**, $\exists y.\varphi[x/y]$ is false, and so $\exists y.\varphi[x/y] \Rightarrow \varphi[x/t]$ is true for *all* t. And so, given that we want to assume extensionality, $\varepsilon x.\varphi$ must be the same if φ is a null property as it would be if φ is universal (cf. the discussions in [Leisenring, pp. 34–35] and [Asser, p. 65]).

Whether we take the detour through η to be truly indicative of the intentions of Hilbert and Bernays as to the meaning of ε or we agree with Leisenring that η is merely a heuristic device, the need to assume $(\varepsilon x.x = x) = (\varepsilon x.x \neq x)$ obviously depends crucially on our assumption of extensionality. So if we find that we are in a position of having to say strange things, we can salve our consciences by putting the blame on supposing that ideal objects can be chosen extensionally–especially since we will see good reason for rejecting this assumption in the intuitionistic case below.

Since we are working with the extensional case for the present, though, we cannot avoid the mentioned equation. But it is not obvious that we can simply adopt it as our sole additional axiom as Hermes did in the two valued case. Indeed, we will not restrict our attention to the extensional classical ε–calculus and the classical ε^*–calculus, because it does not fit naturally with the simple semantics suggested by the remarks above. Perhaps the easiest way to see this is to consider what happens when we try to construct a canonical model for these calculi. As usual, the Lindenbaum algebra is a Boolean algebra, and we can define the interpretations of the predicates and terms in the usual way. The problem comes when we try to define our choice function E. In earlier cases we have adapted to specific needs the approach of defining $E(N)$ to be $[\varepsilon x.\varphi]$ if the formula φ and the variable x represent N (in an sense appropriate for the calculus we are working with at the time), and to be an arbitrary element of N otherwise. Unfortunately, E so defined is not in general a choice function in the current case, because it is possible that both $\langle\varphi, x\rangle$

and $\langle\psi, y\rangle$ represent N and yet $\nvdash \varepsilon x.\varphi = \varepsilon y.\psi$, so E is not even a function.

This problem, of course, has its roots in the fact that we have more than two truth-values. This requires us to modify our definition of what it means for a formula to represent a set. Suppressing the details for the present, we must say that $\langle\varphi, x\rangle$ represents N if $\{\, d \in D \mid \|\varphi[x/d]\| = \bigvee_{d\in D} \|\varphi[x/d]\|\,\} = N$. But then N might be represented by both $\langle\varphi, x\rangle$ and $\langle\psi, y\rangle$ but $\|\exists x.\varphi\| \neq \|\exists y.\psi\|$, and so $\|\forall x.(\varphi \Leftrightarrow \psi)\| \neq 1$. So neither (Ack) nor the $(*)$ rule will guarantee that $\vdash \varepsilon x.\varphi = \varepsilon y.\psi$. We will therefore introduce modified versions of (Ack) and $(*)$ which circumvent this problem. As a result we will be working with four different calculi in this section.

There is another place where our earlier work depends crucially on the fact that we were working with **2** as a lattice of truth-values, namely in our proof of the Soundness Theorem for the classical extensional ε–calculus. Specifically, we can only conclude on the basis of

$$(\dagger) \qquad \|\forall x.(\varphi \Leftrightarrow \psi)\|^{\varrho}_{\mathcal{M}} = 1 \implies [\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = [\varepsilon x.\psi]^{\varrho}_{\mathcal{M}}$$

that

$$\|\forall x.(\varphi \Leftrightarrow \psi) \Rightarrow \varepsilon x.\varphi = \varepsilon x.\psi\|^{\varrho}_{\mathcal{M}} = 1$$

if we assume that $\|\forall x.(\varphi \Leftrightarrow \psi)\|^{\varrho}_{\mathcal{M}} \neq 1$ implies $\|\forall x.(\varphi \Leftrightarrow \psi)\|^{\varrho}_{\mathcal{M}} = 0$, since

$$\|\forall x.(\varphi \Leftrightarrow \psi) \Rightarrow \varepsilon x.\varphi = \varepsilon x.\psi\|^{\varrho}_{\mathcal{M}} = 1 \iff \|\forall x.(\varphi \Leftrightarrow \psi)\|^{\varrho}_{\mathcal{M}} \leqslant \|\varepsilon x.\varphi = \varepsilon x.\psi\|^{\varrho}_{\mathcal{M}}.$$

While $(\dagger)$ continues to hold in the Boolean valued case, it no longer follows that (Ack) is valid. Which is, of course, why we need to work with both the extensional ε-calculus and the ε^*-calculus. The difference between these two calculi is of great interest in the intuitionistic case, since (Ack), given some rather weak subsidiary assumptions, implies excluded middle, while the $(*)$ rule does not. But the relationship is also of some interest in the classical case.

2.2. Back to Work. We begin our investigation of Boolean valued semantics for the ε-calculus with some terminology.

Definition 2. Interpreting the ε-terms in the Boolean valued case is rather more complicated than it was in Chapter 7. If $\mathcal{M}$ is a Boolean valued structure, and presuming everything involved is well-defined, we will say that a join which has the form $\bigvee_{d\in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ is *attained* in $\mathcal{M}$ if there is at least one $d \in D$ such that $\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \bigvee_{d\in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$. For $n \in \omega$, we will say that a Boolean valued structure for $\mathcal{L}'$ is n-join-attained if for all valuations ϱ, $x \in \mathbf{Var}$, and φ with n or fewer occurrences of ε, the join $\bigvee_{d\in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ is attained.

Now we can extend the definitions of $[t]_{\mathcal{M}}^{\varrho} \in D$ and $\|\varphi\|_{\mathcal{M}}^{\varrho} \in B$ from the terms and formulas of $\mathcal{L}'$ to all of **Terms** and **Wffs** by rewriting Definition 1, except that we write $\mathcal{L}$ for $\mathcal{L}'$, and we say that if $\mathcal{M}$ is n-join-attained, then for every φ with n or fewer occurrences of ε,:

$$[\varepsilon x.\varphi]_{\mathcal{M}}^{\varrho} = E\left(\{\, d \in D \mid \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \bigvee_{d\in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} \}\right).$$

If $\mathcal{M}$ is n-join attained for all $n \in \omega$, $\mathcal{M} = \langle D, B, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c}, E\rangle$ is a *quasi-extensional classical Boolean valued structure* for $\mathcal{L}$.

The next step is, of course, to check once again that our machinery continues to mesh properly with our definition of substitution by proving analogues of Theorems

7.1 and 7.2. However, we can use the same proof we used in Chapter 7, with only the most trivial modifications, namely replacing sets of the form $\{ d \in D \mid \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \}$ by sets of the form $\{ d \in D \mid \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \bigvee_{d \in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} \}$ at a few places.

If $\mathcal{M}$ is a quasi-extensional Boolean-valued structure for $\mathcal{L}$, and ϱ is a valuation in $\mathcal{M}$, and $\|\varphi\|_{\mathcal{M}}^{\varrho} = 1$, we say that φ is *valid under* ϱ *in* $\mathcal{M}$, and we write '$\mathcal{M}, \varrho \models_q \varphi$'. If $\mathcal{M}, \varrho \models_q \varphi$ for all ϱ in $\mathcal{M}$, we say that $\mathcal{M}$ is a *model of* φ and write '$\mathcal{M} \models_q \varphi$'. If $\Gamma \subseteq$ **Wffs**, we write '$\mathcal{M} \models_q \Gamma$' if $\mathcal{M} \models_q \beta$ for all $\beta \in \Gamma$. We write '$\Gamma \models_q \varphi$' if $\mathcal{M} \models_q \Gamma$ implies $\mathcal{M} \models_q \varphi$.

Definition 3. We will need the following Boolean valued version of (Ack):

(BAck) $$\forall x[(\exists y.\varphi \Rightarrow \varphi[y/x]) \Leftrightarrow (\exists y.\psi \Rightarrow \psi[y/x])] \Rightarrow \varepsilon y.\varphi = \varepsilon y.\psi.$$

We will also need the Boolean valued version of the $(*)$ rule:

(B$*$) $$\frac{\forall x[(\exists y.\varphi \Rightarrow \varphi[y/x]) \Leftrightarrow (\exists y.\psi \Rightarrow \psi[y/x])]}{\varepsilon y.\varphi = \varepsilon y.\psi}.$$

We define the *B-extensional classical ε-calculus* and the *classical Bε^*-calculus* to be the calculi which result if we replace (Ack) by (BAck) in the extensional classical ε-calculus and $(*)$ by (B$*$) in the classical ε^*-calculus respectively.

We will write '$\Gamma \vdash_{\varepsilon^*} \varphi$' for "$\varphi$ is derivable from Γ in the classical ε^*-calculus", and '$\Gamma \vdash_{\varepsilon^{B*}}$' for the same claim in with respect to the classical Bε^*-calculus.

Theorem 4 (ε^* Soundness Theorem). (1) $\Gamma \vdash_{\varepsilon^*} \varphi \implies \Gamma \models_q \varphi$.

(2) $\Gamma \vdash_{\varepsilon^{B*}} \varphi \implies \Gamma \models_q \varphi$.

Proof. By Theorem 1 and the fact that a quasi-extensional Boolean-valued ε-structure is a Boolean-valued structure, it suffices to show that (ε) and (α) are valid schemes, and that in the quasi-extensional structures $(*)$ is a valid inference rule to complete the proof of (1), and that (B$*$) is valid for (2). We have already noted that the analogue of Lemma 7.4 holds, so (α) follows. To show that (ε) is valid, we must show that $\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} \leqslant \|\varphi[x/\varepsilon x.\varphi]\|^{\varrho}_{\mathcal{M}}$, which follows directly from the fact that our definitions ensure that $\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi[x/\varepsilon x.\varphi]\|^{\varrho}_{\mathcal{M}}$.

To show that $\|\varepsilon x.\varphi = \varepsilon x.\psi\|^{\varrho}_{\mathcal{M}} = 1$, it suffices to show that, for $A = \{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d\in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} \,\}$ and $B = \{\, d \in D \mid \|\psi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d\in D} \|\psi\|^{\varrho(x/d)}_{\mathcal{M}} \,\}$, $A = B$. Now $\|\forall x.(\varphi \Leftrightarrow \psi)\|^{\varrho}_{\mathcal{M}} = 1$ if and only if $\|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \|\beta\|^{\varrho(x/d)}_{\mathcal{M}}$ for all $d \in D$, which clearly implies that $A = B$. So (1) holds. But that $A = B$ also follows from

$$\|(\exists x.\varphi \Rightarrow \varphi[x/y])\|^{\varrho(x/d)}_{\mathcal{M}} = \|(\exists x.\psi \Rightarrow \psi[x/y])\|^{\varrho(x/d)}_{\mathcal{M}} \qquad \text{for all } d \in D.$$

So (2) holds too. $\square$

Unfortunately, we cannot use these quasi-extensional structures to prove a soundness theorem for the extensional or B–extensional calculus because (Ack) is not valid in the quasi-extensional structures. To see this, consider the following

Example 2. Let B be the eight element Boolean algebra. Label the elements of B with the eight possible combinations of three 0's and 1's so that 0=000 and 1=111, and for $a, b \in B$, $a \leqslant b$ if and only if each digit of a is less than or equal to the corresponding digit of b.

Let D be some set, and $f : D \to B$ some *onto* function. Let $\mathcal{L}$ be a language without constants, and with only two unary predicates, $P(x)$ and $R(x)$. Let $R(x)$

be interpreted by $\mathcal{R}_1 : D \to B$ such that

$$\mathcal{R}_1(d) = \begin{cases} 1 & \text{if } d \in \{\, d' \mid f(d') \geqslant 010 \,\} \\ f(d) & \text{otherwise} \end{cases}$$

and let $P(x)$ be interpreted by $\mathcal{R}_2 : D \to B$ such that

$$\mathcal{R}_2(d) = \begin{cases} 1 & \text{if } d \in \{\, d' \mid f(d') \geqslant 110 \,\} \\ f(d) & \text{otherwise.} \end{cases}$$

Let $\mathrm{eq}_{\mathcal{M}}$ be the function such that $\mathrm{eq}_{\mathcal{M}}(d_1, d_2) = 1$ if $d_1 = d_2$, and $\mathrm{eq}_{\mathcal{M}}(d_1, d_2) = 0$ otherwise.

Finally let E be a choice function on $\mathcal{P}(D)$ such that

$$E(\{\, d \in D \mid f(d) \geqslant 010 \,\}) \neq E(\{\, d \in D \mid f(d) \geqslant 110 \,\}).$$

Recalling that f is onto it is easy to see that $\mathcal{M} = \langle D, B, \mathrm{eq}_{\mathcal{M}}, \{\, R_i : i = 1, 2 \,\}, \emptyset, E \rangle$ is a quasi-extensional Boolean-valued structure for $\mathcal{L}$. However,

$$\|\forall x.(P(x) \Leftrightarrow R(x))\| = 010,$$

but

$$[\varepsilon x.P(x)] = E(\{\, d \in D \mid f(d) \geqslant 010\}) \neq E(\{\, d \in D \mid f(d) \geqslant 110\}) = \varepsilon x.R(x),$$

and so $\|\varepsilon x.P(x) = \varepsilon x.R(x)\| = 0$. But then

$$\|\forall x.(P(x) \Leftrightarrow R(x))\| \nleqslant \|\varepsilon x.P(x) = \varepsilon x.R(x)\|,$$

and so (Ack) is not valid in this structure. Since $\forall x.(\varphi \Leftrightarrow \psi)$ implies $\forall x.[(\exists y.\varphi \Rightarrow \varphi[y/x]) \Leftrightarrow (\exists y.\psi \Rightarrow \psi[y/x])]$, this also shows that (BAck) is not valid in this structure.

So we will need to introduce a further refinement in our definitions of our structures. We will say that a quasi-extensional Boolean valued ε-structure $\mathcal{M}$ is *officious* if for all $\varphi, \psi \in$ **Wffs** and every valuation ϱ in $\mathcal{M}$

$$\bigwedge_{d \in D} \|\varphi \Leftrightarrow \psi\|_{\mathcal{M}}^{\varrho(x/d)} \leqslant \mathrm{eq}_{\mathcal{M}}([\varepsilon x.\varphi]_{\mathcal{M}}^{\varrho}, [\varepsilon x.\psi]_{\mathcal{M}}^{\varrho})$$

and we say it is *punctilious* if

$$\bigwedge_{d \in D} \|(\exists y.\varphi \Rightarrow \varphi[y/x]) \Leftrightarrow (\exists y.\psi \Rightarrow \psi[y/x])\|_{\mathcal{M}}^{\varrho(x/d)} \leqslant \mathrm{eq}_{\mathcal{M}}([\varepsilon x.\varphi]_{\mathcal{M}}^{\varrho}, [\varepsilon x.\psi]_{\mathcal{M}}^{\varrho}).$$

(For those who might be wondering about these names: There is a very rough analogy between the necessary condition of for a λ-algebra in a ccc being weakly extensional, i. e. the category *having enough points* and the form the two conditions currently under discussion take if we interpret $=$ by identity. In that case (Ack), for instance, becomes the condition that $[\varepsilon x.\varphi]_{\mathcal{M}}^{\varrho} \neq [\varepsilon x.\psi]_{\mathcal{M}}^{\varrho} \implies \bigwedge_{d \in D} \|\varphi \Leftrightarrow \psi\|_{\mathcal{M}}^{\varrho} = 0$. So the adjectives *punctilious* and *officious* were chosen to suggest using fine points to make subtle discriminations.)

In effect, restricting attention to officious or punctilious structures is simply insisting that (Ack) and (BAck) be satisfied in our structures, since these conditions are simply restatements of these axioms in semantic form. So it is clear what we must do to get a soundness proof for the extensional case.

Definition 4. $\mathcal{M}$ is an *extensional Boolean valued ε-structure* if $\mathcal{M}$ is a punctilious quasi-extensional Boolean valued ε-structure.

We will use '$\models$' to replace '$\models_q$' when we are restricting attention to extensional structures. We will write '$\Sigma \vdash_{ext\varepsilon} \varphi$' for "$\varphi$ is derivable from Σ in the extensional classical ε-calculus," and simply '$\Sigma \vdash_\varepsilon \varphi$' for "$\varphi$ is derivable from Σ in the B-extensional classical ε-calculus," We can use Theorem 4 and the obvious fact that if $\mathcal{M}$ is punctilious it is officious to get

Theorem 5 (Soundness Theorem). (1) $\Sigma \vdash_{ext\varepsilon} \varphi \implies \Sigma \models \varphi$.

(2) $\Sigma \vdash_\varepsilon \varphi \implies \Sigma \models \varphi$. □

To prove completeness we only need to slightly modify the Lindenbaum algebra proof given in §1. For the present we will write '$\vdash$' to stand ambiguously for $\vdash_{\varepsilon B}$. and $\vdash_\varepsilon$. First, we put

$$\varphi \approx \psi \iff \Sigma \vdash \varphi \Leftrightarrow \psi$$

and we write '$\|\varphi\|$' for $\{\, \psi \in \mathbf{Wffs} \mid \varphi \approx \psi \,\}$. Put $B = \{\, \|\varphi\| \mid \varphi \in \mathbf{Wffs} \,\}$, and order B by

$$\|\varphi\| \leqslant \|\psi\| \iff \Sigma \vdash \varphi \Rightarrow \psi.$$

We can show as usual that $\langle B, \leqslant \rangle$ is a Boolean algebra, and that

$$\|\exists x.\varphi\| = \bigvee_{t \in \mathbf{Terms}} \|\varphi[x/t]\|$$

and

$$\|\forall x.\varphi\| = \bigwedge_{t \in \mathbf{Terms}} \|\varphi[x/t]\|\,.$$

Now we define

$$t \sim s \iff \Sigma \vdash t = s$$

and put $[t] = \{\, s \in \mathbf{Terms} \mid t \sim s \,\}$. We will also put

$$[\mathbf{Terms}] = \{\, [t] \mid t \in \mathbf{Terms} \,\}.$$

It is routine to show that we can define, for $P_i \in \mathbf{Pred}_n$,

$$\mathcal{R}_i([t_1], \ldots, [t_n]) = \|P_i t_1 \ldots t_n\| \,,$$

that this is independent of choice of representative for the $[t_i]$, and so

$$\bigvee_{[t]\in[\mathbf{Terms}]} \|\varphi[x/t]\| = \bigvee_{t\in\mathbf{Terms}} \|\varphi[x/t]\|$$

and similarly for $\bigwedge$. By putting $\mathbf{c}(j) = [c_j]$ for $j \in J$, we get a quasi-extensional Boolean valued ε–structure for $\mathcal{L}$, provided we can define an appropriate choice function E. This requires that we modify our definition of what we mean by 'representing'.

Definition 5. Let $\varphi \in$ **Wffs** with free variables among $\{x, z_1, \ldots, z_n\}$, and $t_1, \ldots, t_n \in$ **Terms**. We use the abbreviations $\vec{z} = \langle z_1, \ldots, z_n\rangle$ and $\vec{t} = \langle t_1, \ldots, t_n\rangle$. Let $N \subseteq$ [**Terms**]. We say that the triple $\langle \varphi, x, \vec{t}\rangle$ *represents* N (and that N is *represented* by the triple, and so that N is *representable*) if

$$N = \{\, [s] \in [\mathbf{Terms}] \mid \|\varphi[x/s, \vec{z}/\vec{t}]\| = \bigvee_{[s]\in[\mathbf{Terms}]} \|\varphi[x/s, \vec{z}/\vec{t}]\| \,\}.$$

We need

Lemma 4. *If* $\langle \varphi, x, \vec{t}\rangle$ *and* $\langle \psi, x', \vec{t'}\rangle$ *both represent* $N \subseteq$ [**Terms**], *then* $\Sigma \vdash \varepsilon x.\varphi = \varepsilon x'.\psi$.

Proof. We have $\mathbf{FV}(\varphi) \subseteq \{x, z_1, \ldots, z_n\}$ and $\mathbf{FV}(\psi) \subseteq \{x', z'_1, \ldots, z'_l\}$ for some n and l and we can assume without loss of generality that all these variables are distinct. We can come up with a sequence of $n+l+1$ variables which occur nowhere in φ

or ψ, $\vec{w} = \langle w, w_1, \ldots, w_{l+n}\rangle$ and a sequence of terms $\vec{s} = \langle t_1, \ldots, t_n, t'_1, \ldots, t'_l\rangle$. Now if we take $\varphi' \equiv \varphi[x/w, z_1/w_1, \ldots, z_n/w_n]$ and $\psi' = \psi[x/w, z'_1/w_{n+1}, \ldots, z'_l/w_{n+l}]$ it is easily seen that for any $t \in$ **Terms**

$$\varphi'[w/t, \vec{w}/\vec{s}] \equiv \varphi[x/t, \vec{z}/\vec{t}]$$

and

$$\psi'[w/t, \vec{w}/\vec{s}] \equiv \psi[x/t, \vec{z'}/\vec{t'}\,].$$

So N is represented by both $\langle\varphi', w, \vec{s}\rangle$ and $\langle\psi', w, \vec{s}\rangle$.

Consider the term

$$r \equiv \varepsilon w.\neg[(\exists w.\varphi'[\vec{w}/\vec{s}] \Rightarrow \varphi'[\vec{w}/\vec{s}]) \Leftrightarrow (\exists w.\psi'[\vec{w}/\vec{s}] \Rightarrow \psi'[\vec{w}/\vec{s}])].$$

Now

$$r \in N \iff \|\varphi'[w/r, \vec{w}/\vec{s}]\| = \bigvee_{[s]\in[\mathbf{Terms}]} \|\varphi'[w/s, \vec{w}/\vec{s}]\|$$

and

$$r \in N \iff \|\psi'[w/r, \vec{w}/\vec{s}]\| = \bigvee_{[s]\in[\mathbf{Terms}]} \|\psi'[w/s, \vec{w}/\vec{s}]\|\,,$$

which is to say $\Sigma \vdash \varphi'[w/r, \vec{w}/\vec{s}] \Leftrightarrow \exists w.\varphi'[\vec{w}/\vec{s}]$ if and only if $\Sigma \vdash \psi'[w/r, \vec{w}/\vec{s}] \Leftrightarrow \exists w.\psi'[\vec{w}/\vec{s}]$. It follows by the deduction theorem (twice) and the definition of $\Leftrightarrow$ that

$$\Sigma \vdash (\varphi'[w/r, \vec{w}/\vec{s}] \Leftrightarrow \exists w.\varphi'[\vec{w}/\vec{s}]) \Leftrightarrow (\psi'[w/r, \vec{w}/\vec{s}] \Leftrightarrow \exists w.\psi'[\vec{w}/\vec{s}]),$$

whence

$$\Sigma \vdash (\exists w.\varphi'[\vec{w}/\vec{s}] \Rightarrow \varphi'[w/r, \vec{w}/\vec{s}]) \Leftrightarrow (\exists w.\psi'[\vec{w}/\vec{s}] \Rightarrow \psi'[w/r, \vec{w}/\vec{s}]).$$

So by Lemma 7.11(b) and our choice of r,

$$\Sigma \vdash \forall x.[(\exists w.\varphi'[\vec{w}/\vec{s}] \Rightarrow \varphi'[\vec{w}/\vec{s}]) \Leftrightarrow (\exists w.\psi'[\vec{w}/\vec{s}] \Rightarrow \psi'[\vec{w}/\vec{s}])],$$

and so by $(*)$ or (BAck), as the case may be, we have

$$\varepsilon x.\varphi[\vec{z}/\vec{t}] \equiv \varepsilon x.\varphi'[\vec{w}/\vec{s}] = \varepsilon x.\psi'[\vec{w}/\vec{s}] = \varepsilon x'\psi[\vec{w}/\vec{s}]$$

with the last equality justified by (α). $\square$

So we can define our choice function $E : \mathcal{P}([\mathbf{Terms}]) \to [\mathbf{Terms}]$ by putting

$$E(N) = [\varepsilon x.\varphi[\vec{z}/\vec{t}]]$$

if N is represented by $\langle \varphi, x, \vec{t} \rangle$, and if $N \neq \emptyset$ is not representable, we can let $E(N)$ be an arbitrary member of N. The reader will notice that we have defined our current structures so that E need not assign anything to $\emptyset$ (which is obviously okay since we require the joins used to interpret existential quantifiers to be attained), so this definition is sufficient.

We will use the map $\pi(x) = [x]$ which takes **Var** into [**Terms**]. Let $i : B \to B'$ be an isomorphism of the type guaranteed to exist by Proposition 1. We thus have a structure

$$\mathcal{M} = \langle [\mathbf{Terms}], B', \{\, i \circ \mathcal{R}_i : i \in I \,\}, i \circ \|\cdot = \cdot\|, \mathbf{c}, E \rangle$$

and it is straightforward to show that for each φ

$$\|\varphi\|^{\pi}_{\mathcal{M}} = i(\|\varphi\|).$$

π is the canonical valuation in $\mathcal{M}$. It is easy to show that $\mathcal{M}$ is a quasi-extensional Boolean valued ε–structure, and by Theorem 4 if we have (BAck) rather than (B$*$) it is an extensional Boolean valued structure.

Theorem 6 (Completeness). *Let Σ be a set of sentences of $\mathcal{L}$. The following are equivalent.*

(1) $\Sigma \vdash_{\varepsilon} \varphi$

(2) $\Sigma \models \varphi$

(3) $\|\varphi\|^{\pi}_{\mathcal{M}} = 1$ *on the canonical valuation.*

The following are also equivalent.

(1′) $\Sigma \vdash_{\varepsilon B*} \varphi$

(2′) $\Sigma \models_{q} \varphi$

(3′) $\|\varphi\|^{\pi}_{\mathcal{M}} = 1$ *on the canonical valuation.*

Proof. The proof is the same for both parts. (1) $\Longleftrightarrow$ (3) is clear by our construction. (2) $\Longrightarrow$ (3) is trivial, so (2) $\Longrightarrow$ (1). (1) $\Longrightarrow$ (2) is Soundness. □

In §1 we went on from this point to use the Ultrafilter Theorem to reduce the Boolean valued case to the **2**–valued case. Unfortunately, this is not an option in the ε–calculus case.

Of course, the Ultrafilter Theorem still tells us that the Lindenbaum algebra for Σ contains an ultrafilter U, and if we disregard E we have a Boolean valued structure for the first order language $\mathcal{L}'$. So we can get a **2**–valued model for Σ by the same construction as in §1. But what happens to E if we try to treat the ε–terms in the same way as other terms?

If we put $\tilde{t} = \{\, t' \mid t = t' \in U \,\}$, and $\mathcal{T} = \{\, \tilde{t} \mid t \in \mathbf{Terms} \,\}$, we will induce a map $\tilde{}: [\mathbf{Terms}] \to \mathcal{T}$, since if $\vdash t = t'$, $\|t = t'\| = 1 \in U$. Put $\tilde{N} = \{\, \tilde{t} \mid t \in N \,\}$ for $N \subseteq [\mathbf{Terms}]$. Unfortunately it is not in general true that if $\langle \varphi, x, \vec{t} \rangle$ represents

$N \subseteq$ [**Terms**] it will represent $\check{N} \subseteq \mathcal{T}$, or indeed that $\check{N}$ will be representable at all. And there may well be representable subsets of $\mathcal{T}$ that are not the $\check{}$-images of representable subsets of [**Terms**]. In short, the result of this is not in general going to be an ε-structure, since E will often turn out not to be a function.

Now we could simply carry through the construction outlined at the end of §1, then rewrite our earlier definition of E and so define a *new* choice function E' in the resulting **2**-valued structure. We will see below that the result would be a model of Σ, (Ack) and the formula $(\varepsilon x.x = x) = (\varepsilon x.x \neq x)$. However, factoring by an ultrafilter is probably more interesting if we *ignore* E.

While factoring by an ultrafilter does not give us an ε-structure, it nonetheless does give us a **2**-valued structure for $\mathcal{L}$—in applying this procedure we simply "forget" the E part of our structure. Note also that every ultrafilter in the Lindenbaum algebra $\langle B, \leqslant\rangle$ is *perfect* since $\|\exists x.\varphi\| = \|\varphi[x/\varepsilon x.\varphi]\|$, so by Theorem 2 we have a **2**-valued model of Σ.

To see why this is interesting, assume Σ and φ are ε-free. We have just constructed a canonical model for Σ, so we can prove $\Sigma \vdash_\varepsilon \varphi \iff \|\varphi\|^{\pi}_{\mathcal{M}} = 1$. In Chapter 7 we gave semantic proofs of the ε-theorems, mostly because they are much shorter than the traditional syntactic proofs. But we could (with a lot more effort than is warranted by the small point to be made here) give a purely syntactical proof of the ε-theorems for the classical B-extensional ε-calculus. (Curious readers should be able to construct such a proof on the model of the syntactic proofs for the classical extensional ε-calculus in Chapter 3 of [Leisenring].) So we have

$\Sigma \vdash_{PC} \varphi \iff \|\varphi\|^{\pi}_{\mathcal{M}} = 1$, and from this we can get a proof of Theorem 3. What is interesting is that we have this proof for the cases of *both countable and uncountable languages* $\mathcal{L}$, since we have been able to apply the ultrafilter theorem directly rather than having to appeal to the Rasiowa–Sikorski Theorem.

In one sense it is not hard to see why the detour through the ε–calculus does this for us. Consider again the two examples for §1. No case like Example 1 can occur since the ε–scheme guarantees that every ultrafilter in the Lindenbaum algebra is perfect. And none like Example 2 can occur since that example turns crucially on the fact that $\mathcal{L}'$ has only countably many terms, but uncountably many predicates. The addition of ε to a language clearly guarantees that if there are uncountably many predicates there will be uncountably many terms as well. But a proper explanation of what is going on here would require an algebraic characterization of the move from the Lindenbaum algebra for Σ in $\mathcal{L}'$ to the Lindenbaum algebra for Σ in $\mathcal{L}$. This, as far as I have been able to determine, is an interesting, messy and still open problem.

Now in Chapter 7 we already presented a complete semantics for the classical extensional ε–calculus. We conclude this chapter by using the two systems of semantics, together with the fact that in a **2**–valued $\mathcal{M}$

$$\{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 1 \,\} = \{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} \,\},$$

to make a few observations.

Let us call the following sentence '(Her)', since it was adopted as an axiom by

Hermes:

$$\text{(Her)} \qquad (\varepsilon x.x = x) = (\varepsilon x.x \neq x).$$

It is not hard to see that both (Ack) and (Her) are derivable from (BAck). Similarly, if (B$*$) is valid, then (*) is valid and (Her) is derivable.

Now, if we restrict attention to the **2**–valued case, (Ack) and (Her) together imply (BAck) (and so the corresponding remark holds for the ($*$) case). For suppose

$$\|\forall x.[(\exists y.\varphi \Rightarrow \varphi[y/x]) \Leftrightarrow (\exists y.\psi \Rightarrow \psi[y/x])]\|^{\varrho}_{\mathcal{M}} = 1$$

(if it is 0, then (BAck) is trivially satisfied). Then for each $d \in D$

$$\|\exists y.\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} \iff \|\exists y.\psi\|^{\varrho}_{\mathcal{M}} = \|\psi\|^{\varrho(x/d)}_{\mathcal{M}}$$

and

$$\|\exists y.\varphi\|^{\varrho}_{\mathcal{M}} \neq \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} \iff \|\exists y.\psi\|^{\varrho}_{\mathcal{M}} \neq \|\psi\|^{\varrho(x/d)}_{\mathcal{M}}.$$

Now suppose $\|\exists y.\varphi\|^{\varrho}_{\mathcal{M}} = \|\exists y.\psi\|^{\varrho}_{\mathcal{M}}$. Then for each $d \in D$, $\|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \|\psi\|^{\varrho(x/d)}_{\mathcal{M}}$, so $\|\forall x.(\varphi \Leftrightarrow \psi)\|^{\varrho}_{\mathcal{M}} = 1$, so $\|\varepsilon x.\varphi = \varepsilon x.\psi\|^{\varrho}_{\mathcal{M}} = 1$ by (Ack) or ($*$). If on the other hand $\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} \neq \|\exists x.\psi\|^{\varrho}_{\mathcal{M}}$, then one is 0 and the other is 1. So say $\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} = 0$. Then $\|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 0$ for all $d \in D$, so $\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}}$ for all d. But then $\|\psi\|^{\varrho(x/d)}_{\mathcal{M}} = 1$ for all d. So $\|\forall x.(\varphi \Leftrightarrow x \neq x)\|^{\varrho}_{\mathcal{M}} = 1$, so $\|\varepsilon x.\varphi = \varepsilon x.x \neq x\|^{\varrho}_{\mathcal{M}} = 1$ by (Ack) or ($*$), and $\|\forall x.(\psi \Leftrightarrow x = x)\|^{\varrho}_{\mathcal{M}} = 1$ so $\|\varepsilon x.\psi = \varepsilon x.x = x\|^{\varrho}_{\mathcal{M}} = 1$, again by (Ack) or ($*$). So $\|\varepsilon x.\varphi = \varepsilon x.\psi\|^{\varrho}_{\mathcal{M}} = 1$ since (Her) is valid. But then it follows from the Completeness Theorem 7.10:

Proposition 2. *In the classical ε-calculus (BAck) is equivalent to the conjunction of (Ack) and (Her).* □

Of course, we could also have proved this syntactically. This only holds in the classical case, so while we could have simply added (Her) as an axiom in this chapter the proofs we wound up with would then have been less useful as examples for our treatment of the intuitionistic case below.

If we restrict attention to **2**–valued structures, the difference between (B$*$) and (BAck) (and the difference between ($*$) and (Ack)) disappears, as we have already mentioned. So there is no difference between quasi-extensional and extensional **2**-valued ε–structures. However, as we have seen, there are quasi-extensional Boolean valued structures which are not extensional, hence formulas valid in all Boolean valued extensional ε–structures which are not valid in all Boolean valued quasi–extensional ε–structures. We have shown that there are enough **2**–valued extensional structures—hence enough **2**–valued quasi-extensional structures—for us to be able to prove completeness for the extensional ε–calculus. But we also have a complete Boolean valued semantics for the extensional ε–calculus, so a formula φ is valid in all Boolean valued extensional structures if and only if φ is valid in all **2**–valued extensional ε–structures if and only if φ is valid in all **2**–valued quasi–extensional ε–structures. So we can conclude

Proposition 3. *The ε^*–calculus is not complete for* **2**–*valued quasi-extensional ε-semantics.* □

There does not seem to be any natural **2**–valued semantics of the sort we have been considering for the ε^*-calculus. However, in Chapter 12 §3 I suggest some reason to think that the difference between extensionality and quasi-extensionality is best understood in modal terms.

CHAPTER X

FIRST ORDER INTUITIONISTIC SEMANTICS

This chapter is intended to play a role for intuitionistic logic similar to that played by Chapter 6 for classical logic. The material is perhaps less well-known than the material for classical logic, but it is available in a wide variety of sources and so again proofs are often omitted. We will also introduce some facts about intuitionistic logic that will be relevant to our discussion of the effects of adding ε and τ to intuitionistic logic. Unlike the classical case, the addition of ε or τ to intuitionistic predicate calculus is non-conservative, so we will list some schemes valid in classical logic but not valid in intuitionistic logic because it is a natural question whether ε (or some other operator) makes each of them valid.

Once again, for $\mathcal{L}$ a language as defined in Chapter 5, we will work with the language for the first order predicate calculus with equality we get by taking the subset of the well-formed expressions $\mathcal{L}' = \{ M \mid M \text{ is } \varepsilon\text{–free} \}$. We again assume that the predicate symbols and constants of $\mathcal{L}'$ are given in the form of indexed sets $\{ P_i \mid i \in I \}$ and $\{ c_j \mid j \in J \}$, and for each $i \in I$ we denote the arity of P_i by $\zeta(i)$.

An *intuitionistic* $\mathcal{L}'$ *structure* is an ordered quintuple $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c} \rangle$ where

(1) $D \neq \emptyset$ is a set

(2) Ω is a complete Heyting Algebra

(3) $\mathcal{R} : I \rightarrow \{ f \mid$ for some positive integer n, $f : D^n \rightarrow \Omega \}$ is such that for each $i \in I$, $\mathcal{R}_i : D^{\zeta(i)} \rightarrow \Omega$.

(4) $\mathrm{eq}_{\mathcal{M}} : D \times D \rightarrow \Omega$ is such that for all $d_1, d_2, d_3, d_4, \ldots, d_{2\zeta(i)} \in D$ and $\mathcal{R}_i$,

   (i) $\mathrm{eq}_{\mathcal{M}}(d_1, d_1) = 1$

   (ii) $\mathrm{eq}_{\mathcal{M}}(d_1, d_2) = \mathrm{eq}_{\mathcal{M}}(d_2, d_1)$

   (iii) $\mathrm{eq}_{\mathcal{M}}(d_1, d_2) \wedge \mathrm{eq}_{\mathcal{M}}(d_2, d_3) \leqslant \mathrm{eq}_{\mathcal{M}}(d_1, d_2)$

   (iv) $\mathrm{eq}_{\mathcal{M}}(d_1, d_{\zeta(i)+1}) \wedge \cdots \wedge \mathrm{eq}_{\mathcal{M}}(d_{\zeta(i)}, d_{2\zeta(i)}) \wedge \mathcal{R}_i(d_1, \ldots, d_{\zeta(i)})$
   $\leqslant R_i(d_{\zeta(i)+1}, \ldots, d_{2\zeta(i)})$.

(5) $\mathbf{c} : J \rightarrow D$.

As usual, we define a valuation in $\mathcal{M}$ to be a map $\varrho : \mathbf{Var} \rightarrow D$, and we define the valuation $\varrho(v_n/d) : \mathbf{Var} \rightarrow D$ by

$$\varrho(v_n/d) = \begin{cases} \varrho(v_i) & \text{for } i \neq n \\ d & \text{if } i = n. \end{cases}$$

Definition 1. We define for each term t and each formula φ of $\mathcal{L}'$ the *interpretations,* $[t]^{\varrho}_{\mathcal{M}} \in D$ *and* $\|\varphi\|^{\varrho}_{\mathcal{M}} \in \Omega$, *of* t *and* φ *in* $\mathcal{M}$ *under* ϱ recursively as follows.

$$[x]^{\varrho}_{\mathcal{M}} = \varrho(x) \qquad \text{for } x \in \mathbf{Var}$$

$$
\begin{aligned}
[c_i]^{\varrho}_{\mathcal{M}} &= \mathbf{c}(i) \qquad\qquad \text{for } c_i \in \mathbf{Con} \\
\|P_i t_1 \ldots t_{\zeta(i)}\|^{\varrho}_{\mathcal{M}} &= \mathcal{R}_i\left(\langle [t_1]^{\varrho}_{\mathcal{M}}, \ldots, [t_{\zeta(i)}]^{\varrho}_{\mathcal{M}} \rangle\right) \\
\|t_1 = t_2\|^{\varrho}_{\mathcal{M}} &= \mathrm{eq}_{\mathcal{M}}([t_1]^{\varrho}_{\mathcal{M}}, [t_2]^{\varrho}_{\mathcal{M}}) \\
\|\neg\varphi\|^{\varrho}_{\mathcal{M}} &= (\|\varphi\|^{\varrho}_{\mathcal{M}})^* \\
\|\varphi \wedge \psi\|^{\varrho}_{\mathcal{M}} &= \|\varphi\|^{\varrho}_{\mathcal{M}} \wedge \|\psi\|^{\varrho}_{\mathcal{M}} \\
\|\varphi \vee \psi\|^{\varrho}_{\mathcal{M}} &= \|\varphi\|^{\varrho}_{\mathcal{M}} \vee \|\psi\|^{\varrho}_{\mathcal{M}} \\
\|\varphi \Rightarrow \psi\|^{\varrho}_{\mathcal{M}} &= \|\varphi\|^{\varrho}_{\mathcal{M}} \Rightarrow \|\psi\|^{\varrho}_{\mathcal{M}} \\
\|\exists x.\varphi\|^{\varrho}_{\mathcal{M}} &= \bigvee_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} \\
\|\forall x.\varphi\|^{\varrho}_{\mathcal{M}} &= \bigwedge_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} .
\end{aligned}
$$

The symbols *, $\Rightarrow$, $\vee$, and $\wedge$ or the right denote the operations in Ω. This is pretty close to Definition 6.1. We have changed the clause for $=$ for the same reason given for the corresponding change in Chapter 9, and we have eliminated the parenthetical remark in the clause for $\Rightarrow$. The second change is of course necessitated by the fact that it is not in general true in Heyting algebras that $p \Rightarrow q = p^* \vee q$.

If $\|\varphi\|^{\varrho}_{\mathcal{M}} = 1$ for some $\varrho : \mathbf{Var} \to D$, we say that φ is *satisfiable* (in $\mathcal{M}$), and that it is *satisfied by* ϱ. If $\|\varphi\|^{\varrho}_{\mathcal{M}} = 1$ for all $\varrho : \mathbf{Var} \to D$, we say that $\mathcal{M}$ is an (intuitionistic) *model* for φ, and write '$\mathcal{M} \models_{IC} \varphi$'. If $\mathcal{M} \models_{IC} \varphi$ for all intuitionistic $\mathcal{L}'$ structures $\mathcal{M}$, we say that φ is *valid in the intuitionistic first order predicate calculus*, and we write '$\models_{IC} \varphi$'. If Σ is a set of formulas, we say that Σ is (intuitionistically) *satisfiable* if there is some ϱ in some $\mathcal{M}$ that satisfies every member of Σ. We say that $\mathcal{M}$ is an (intuitionistic) model of Σ if $\mathcal{M}$ is a model of every formula in Σ, and we write '$\mathcal{M} \models_{IC} \Sigma$' for this. We write '$\Sigma \models_{IC} \varphi$' if

$$\mathcal{M} \models_{IC} \Sigma \implies \mathcal{M} \models_{IC} \varphi.$$

We write '$\Sigma \vdash_{IC} \varphi$' if φ is derivable from Σ in the intuitionistic first order predicate calculus with identity.

The proof of the Soundness Theorem is, as usual, straightforward.

Theorem 1 (Soundness Theorem for the Intuitionistic Predicate Calculus).

$$\Sigma \vdash_{IC} \varphi \implies \Sigma \models_{IC} \varphi \quad \square$$

It is easiest to prove completeness by using the Lindenbaum algebra construction to give a canonical model proof of the Completeness Theorem. So if Σ is a set of *sentences* of $\mathcal{L}'$, we define for any $\varphi, \psi \in \mathbf{Wffs}$,

$$\varphi \approx \psi \iff \Sigma \vdash \varphi \Leftrightarrow \psi.$$

This is obviously an equivalence relation. We write '$\|\varphi\|_\Sigma$' for $\{\, \psi \in \mathbf{Wffs} \mid \varphi \approx \psi \,\}$, or we just write '$\|\varphi\|$' if it is clear which set Σ we have in mind. We put $\Omega = \{\, \|\varphi\| \mid \varphi \in \mathbf{Wffs} \,\}$, and order Ω by

$$\|\varphi\| \leqslant \|\psi\| \iff \Sigma \vdash \varphi \Rightarrow \psi.$$

This is evidently a sound definition. The proof that Ω is a Heyting algebra is the same as the proof that B was a Boolean algebra in Chapter 9, except that since $\varphi \vee \neg\varphi$ is not an intuitionistically valid schema we can only show that $\|\neg\varphi\|$ is a pseudo-complement for $\|\varphi\|$, not that it is a complement. We can use the proof of

Lemma 9.2, since the required facts about the classical predicate calculus hold also in the intuitionistic case, to get

Lemma 1. *For* $\varphi \in \mathbf{Wffs}$, $x \in \mathbf{Var}$, $t \in \mathbf{Terms}$,

$$\|\exists x.\varphi\| = \bigvee_{t \in \mathbf{Terms}} \|\varphi[x/t]\|$$

and

$$\|\forall x.\varphi\| = \bigwedge_{t \in \mathbf{Terms}} \|\varphi[x/t]\| . \quad \square$$

Ω, so constructed, is the Lindenbaum algebra of Σ. We now have almost all the elements in place to provide a Heyting Algebra valued structure for $\mathcal{L}'$. We can, for $\mathcal{P}_i \in \mathbf{Pred}_n$, put

$$\mathcal{R}_i(t_1, \ldots, t_n) = \|P_i t_1 \ldots t_n\| ,$$

for constants c_i put $\mathbf{c}(i) = c_i$, and for $\mathrm{eq}_{\mathcal{M}}$ we put

$$\mathrm{eq}_{\mathcal{M}}(t_1, t_2) = \|t_1 = t_2\| .$$

However, we have defined Heyting Algebra valued models as taking *complete* Heyting algebras as truth value lattices, and the Lindenbaum algebra need not be complete, so we need to appeal to

Proposition 1. *Every Heyting algebra* Ω *is isomorphic to a subalgebra of a complete Heyting algebra* Ω', *and the injection* $i : \Omega \to \Omega'$ *preserves all joins and meets.* $\square$

For a proof, see [Rasiowa and Sikorski, pp. 139–40]. There are in general many such completions we could appeal to, so we will henceforth assume that when we appeal to Proposition 1 we have selected a particular i and Ω'. We therefore can get a Heyting algebra valued structure $\mathcal{M} = \langle \mathbf{Terms}, \Omega', \{ i \circ \mathcal{R}_i : i \in I \}, i \circ \mathrm{eq}_{\mathcal{M}}, \mathbf{c} \rangle$, and the claims made above for Ω will hold for Ω' also. Obviously we can take the identity map on **Var** as a map $\pi : \mathbf{Var} \to \mathbf{Terms}$, and so as a valuation in $\mathcal{M}$. We call this the canonical valuation of Σ.

It is straightforward to prove

Lemma 2. $\|\varphi\|^{\pi}_{\mathcal{M}} = 1 \iff \Sigma \vdash \varphi$. □

This puts us in a position to prove

Theorem 2. *Let Σ be a set of sentences of $\mathcal{L}'$. Then the following claims are equivalent.*

(1) *Σ is consistent.*

(2) *There is an intuitionistic model for Σ.*

(3) *Ω is non-degenerate.* □

and also

Theorem 3. *The following claims are equivalent.*

(1) $\Sigma \vdash \varphi$.

(2) $\Sigma \models \varphi$.

(3) $\|\varphi\|^{\pi}_{\mathcal{M}} = 1$ *in the canonical valuation of Σ.* □

Now if we restrict our attention to propositional logic, we can in fact add another clause to Theorem 2 since Σ is consistent in intuitionistic logic if and only if Σ is consistent in classical logic. That is, if we are concerned only with propositional logic, we can add the law of excluded middle to an intuitionistic theory without fear of generating an inconsistency. This is no longer the case in predicate logic.

To illustrate this, we will state a few facts. First, we define a theory in $\mathcal{L}'$ to be *maximally intuitionistically consistent* if it is consistent and every extension of it is equivalent to it. We will say an intuitionistic structure $\mathcal{M}$ is *adequate* for a set Σ if $\mathcal{M} \models \varphi \iff \Sigma \vdash \varphi$. Then it can be shown that:

Proposition 2. (1) *Σ is maximally intuitionistically consistent $\iff$ $\Sigma \vdash_{IC} \alpha$ or $\Sigma \vdash_{IC} \neg\alpha$ for all closed formulas α.*

(2) *If Σ is maximal, then every **2**–valued model for Σ is adequate.*

(3) *Every intuitionistically consistent theory can be extended to a maximally intuitionistically consistent theory.*

(4) *If Σ has an adequate **2**–valued model, then Σ is a maximally consistent theory.*

However, suppose $\mathcal{L}'$ has only the one place predicate P, and no constants. Suppose $e : \mathbb{N} \to \mathbb{Q}$ is an enumeration of the rational numbers. Consider the structure we get by taking: $D = \mathbb{N}$, the set of natural numbers; Ω to be the open sets of the real numbers (so $1 = \mathbb{R}$ and $0 = \emptyset$); and $\mathcal{R}_1$ is a function such that for $n \in \mathbb{N}$, $\mathcal{R}_1(n) = \{\, r \in \mathbb{R} \mid r \neq e(n) \,\}$. Then in this structure $\|\forall x.(P(x) \vee \neg P(x))\|_{\mathcal{M}} = \emptyset = 0$, and so $\neg\forall x.(P(x) \vee \neg P(x))$ is satisfied in this structure. Which tells us that if we put $\Sigma = \{\neg\forall x.(P(x) \vee \neg P(x))\}$, Σ is intuitionistically consistent, so by (3) it has a maximally consistent intuitionistic extension. And since $\forall x.(P(x) \vee \neg P(x))$ is classically valid, there can be no **2**–valued model for Σ. So

(5) *There is a maximally intuitionistically consistent theory which has no* **2**-*valued model.*

(6) *There is a set of formulas* Σ *such that* Σ *is consistent in the intuitionistic predicate calculus but which is inconsistent in the classical predicate calculus.* □

A consistent theory with axioms Σ is *prime* if for any *closed* formulas α and β, if $\Sigma \vdash \alpha \vee \beta$ then either $\Sigma \vdash \alpha$ or $\Sigma \vdash \beta$. A theory is *strongly prime* if the same condition holds for arbitrary formulas, not just sentences. The theory is *witnessed* if, for every φ and x, $\Sigma \vdash \exists x.\varphi$ implies $\Sigma \vdash \varphi[x/t]$ for some term t. A theory is *constructive* if it is both strongly prime and witnessed.

Proposition 3. *Whether we are talking about intuitionistic or classical theories,*

(1) *Every maximal consistent theory is prime.*

(2) *Every consistent theory can be extended to a prime theory.*

However, it is well-known that

(3) *The pure intuitionistic predicate calculus is constructive.*

It is also easily seen to be non-maximal. Since (P11) makes it easy to see that primeness implies maximality in the classical case,

(4) *While every prime classical theory is a maximal consistent theory, there are prime intuitionistic theories which are not maximally intuitionistically complete theories.*

We note that

(5) *There are (intuitionistic and classical) theories which are witnessed, but which are not prime.*

Now, every strongly prime classical theory is witnessed. But

(6) *There are strongly prime intuitionistic theories which are not witnessed.* □

Finally, we list some of the formulas which are both valid in the classical predicate calculus and not intuitionistically valid.

Proposition 4. *The following schemes are classically valid, but are not intuitionistically valid.*

(1) $\alpha \vee \neg\alpha$.

(2) $\neg\neg\alpha \Rightarrow \alpha$.

(3) $(\alpha \Rightarrow \beta) \Rightarrow (\neg\alpha \vee \beta)$.

(4) $\neg(\alpha \wedge \beta) \Rightarrow (\neg\alpha \vee \neg\beta)$.

(5) $\neg\neg\alpha \vee \neg\alpha$.

(6) $\neg\neg\alpha \vee \neg\neg\beta \Leftrightarrow \neg\neg(\alpha \vee \beta)$.

(7) $(\alpha \Rightarrow \beta) \vee (\beta \Rightarrow \alpha)$. □

Lemma 3. *The following equivalences are provable in intuitionistic propositional logic.*

(a) (1) $\Longleftrightarrow$ (2) $\Longleftrightarrow$ (3).

(b) (4) $\Longleftrightarrow$ (5) $\Longleftrightarrow$ (6).

Proof. For (a), suppose $\alpha \vee \neg\alpha$ is a valid scheme and assume $\neg\neg\alpha$.

$$(\alpha \Rightarrow \alpha) \Rightarrow ((\neg\alpha \Rightarrow \alpha) \Rightarrow (\alpha \vee \neg\alpha) \Rightarrow \alpha))$$

is an instance of the axiom (P8), and from (P10) and the fact that $\vdash \neg\neg\neg\alpha \Leftrightarrow \neg\alpha$ we know $\neg\neg\alpha \vdash \neg\alpha \Rightarrow \alpha$. Finally, $(\alpha \Rightarrow \alpha)$ is intuitionistically provable, so a few applications of (MP) gives us $\vdash \alpha$. Conversely, $\neg(\neg\neg\alpha \wedge \neg\alpha)$ and $\neg(\neg\neg\alpha \wedge \neg\alpha) \Rightarrow \neg\neg(\neg\alpha \vee \alpha)$ are valid intuitionistic schemes, so (1) follows by (MP) and (2). That (3) $\Longrightarrow$ (1) follows from the fact that $\vdash \alpha \Rightarrow \alpha$. Finally, since $\alpha \Rightarrow \beta, \alpha \vdash \beta$, $\alpha \Rightarrow \beta, \alpha \vee \neg\alpha \vdash \beta \vee \neg\alpha$, so (1) $\Longrightarrow$ (3).

For (b), if $\vdash \neg(\alpha \wedge \beta) \Rightarrow \neg\alpha \vee \neg\beta$ for all α and β, then $\vdash \neg(\alpha \wedge \neg\alpha) \Rightarrow (\neg\alpha \vee \neg\neg\alpha)$. So, since $\neg(\alpha \wedge \neg\alpha)$ is intuitionistically provable, $\vdash \neg\alpha \vee \neg\neg\alpha$.

To prove the converse, notice first that

$$\neg(\alpha \wedge \beta) \vdash \neg(\neg\neg\alpha \wedge \beta),$$

so

$$\neg(\alpha \wedge \beta), \neg\neg\alpha, \beta \vdash [\neg(\neg\neg\alpha \wedge \beta) \wedge (\neg\neg\alpha \wedge \beta)]$$

but $(\psi \Rightarrow (\varphi \wedge \neg\varphi)) \Rightarrow (\psi \Rightarrow \neg\beta)$ is a valid principle of intuitionistic logic, so

$$\neg(\alpha \wedge \beta), \neg\neg\alpha \vdash \neg\beta.$$

Also, obviously, $\neg(\alpha \wedge \beta), \neg\alpha \vdash \neg\alpha$. Thus

$$\begin{aligned}\neg(\alpha \wedge \beta) &\vdash \neg(\alpha \wedge \beta) \wedge (\neg\alpha \vee \neg\neg\alpha)\\ &\vdash [\neg(\alpha \wedge \beta) \wedge \neg\alpha] \vee [\neg(\alpha \wedge \beta) \wedge \neg\neg\alpha]\\ &\vdash \neg\alpha \vee \neg\beta.\end{aligned}$$

So (4) and (5) are equivalent.

To show (5) $\Longrightarrow$ (6), we first note that it is intuitionistically provable that $\neg\neg(\alpha \vee \beta) \Leftrightarrow \neg(\neg\alpha \wedge \neg\beta)$. So by (4) $\vdash \neg\neg(\alpha \vee \beta) \Leftrightarrow \neg\neg\alpha \vee \neg\neg\beta$. Finally, while the law of excluded middle is of course not provable in intuitionistic logic, we do have $\vdash \neg\neg(\alpha \vee \neg\alpha)$, so by (6) $\vdash \neg\neg\alpha \vee \neg\neg\neg\alpha$, and since $\vdash \neg\alpha \Leftrightarrow \neg\neg\neg\alpha$, $\vdash \neg\neg\alpha \vee \neg\alpha$, i. e. (6) $\Longrightarrow$ (5). So (b) is proved. $\square$

Lemma 4. *In intuitionistic propositional logic* (2) $\Longrightarrow$ (7) *and* (7) $\Longrightarrow$ (5), *but in each case the converses fail.*

Proof. It is intuitionistically provable that $\neg(\alpha \Rightarrow \beta) \Leftrightarrow \neg\neg\alpha \wedge \neg\beta$, so $\neg(\alpha \Rightarrow \beta) \vdash \neg\neg\alpha$, so by (2), $\neg(\alpha \Rightarrow \beta), \beta \vdash \alpha$, and by the deduction theorem $\neg(\alpha \Rightarrow \beta) \vdash \beta \Rightarrow \alpha$, and the result follows since (2) implies (1) by Proposition 4, and so we have $\vdash (\alpha \Rightarrow \beta) \vee \neg(\alpha \Rightarrow \beta)$.

Next, note that if $\vdash (\alpha \Rightarrow \neg\alpha) \vee (\neg\alpha \Rightarrow \alpha)$, then $\vdash [\alpha \Rightarrow (\alpha \wedge \neg\alpha)] \vee [\neg\alpha \Rightarrow (\alpha \wedge \neg\alpha)]$, and all formulas of form $\varphi \Rightarrow (\varphi \wedge \neg\varphi) \Rightarrow \neg\varphi$ are intuitionistically provable, so $\vdash \neg\alpha \vee \neg\neg\alpha$.

It is easy to construct counter-examples which show that the converses fail using the following facts. First, a linearly ordered set with both a top and bottom element is a Heyting algebra in which $\alpha \Rightarrow \beta = 1 \iff \alpha \leqslant \beta$, and, for $\beta < 1$, $\alpha \Rightarrow \beta = \beta \iff \beta < \alpha$. So for $\alpha \neq 0$, $\alpha^* = 0$, so if also $\alpha \neq 1$ we have $\alpha^* \vee \alpha = \alpha \neq 1$. But in a chain $(\alpha \Rightarrow \beta) \vee (\beta \Rightarrow \alpha) = 1$, since either $a \leqslant \beta$ or $\beta \leqslant \alpha$. So if we take as our lattice of truth-values any chain with more than two elements, we can construct the required counter-example. Secondly,[1] start with the four element Boolean algebra,

[1] John Bell found this simple example.

write d for the top element, a for the bottom, and b and c for the two non-comparable elements. Now add new top and bottom elements 0 and 1. Call the resulting lattice L. It is not hard to check that $0^* = 1$ and, for $0 \neq \alpha \in L$, $\alpha^* = 0$, so $\alpha^* \vee \alpha^{**} = 1$. But $(b \Rightarrow c) = c$ and $(c \Rightarrow b) = b$, so $(b \Rightarrow c) \vee (c \Rightarrow b) = d \neq 1$. □

Proposition 5. *Assuming x does not occur free in φ, the following schemes are classically valid but are not intuitionistically valid.*

(1) $\neg\forall x.\psi \Rightarrow \exists x.\neg\psi$.

(2) $(\forall x.\psi \Rightarrow \varphi) \Rightarrow \exists x.(\psi \Rightarrow \varphi)$.

(3) $\exists x.(\psi \Rightarrow \forall x.\psi)$.

(4) $(\varphi \Rightarrow \exists x.\psi) \Rightarrow \exists x.(\varphi \Rightarrow \psi)$.

(5) $\exists x.(\exists x.\psi \Rightarrow \psi)$.

(6) $\forall x.(\varphi \vee \psi) \Rightarrow (\varphi \vee \forall x.\psi)$. □

Lemma 5. *The following equivalences are provable in the first order intuitionistic predicate calculus.*

(a) (2) $\Longleftrightarrow$ (3)

(b) (4) $\Longleftrightarrow$ (5)

Proof. Since x is not free in $\forall x.\psi$,

$$(\forall x.\psi \Rightarrow \forall x.\psi) \Rightarrow \exists x.(\psi \Rightarrow \forall x.\psi)$$

is an instance of (2), and $\forall x.\psi \Rightarrow \forall x.\psi$ is obviously provable. Conversely, if both $\exists x.(\psi \Rightarrow \forall x.\psi)$ and $\forall x.\psi \Rightarrow \varphi$ are provable, $\exists x.(\psi \Rightarrow \varphi)$ is provable by the transitivity of provability. So the result follows by the deduction theorem. This proves (a).

As for (b), first note that $(\exists x.\psi \Rightarrow \exists x.\psi) \Rightarrow \exists x.(\exists x.\psi \Rightarrow \psi)$ is an instance of (4). Conversely, $\exists x.(\exists x.\psi \Rightarrow \psi)$ and $\varphi \Rightarrow \exists x.\psi$ together imply $\exists x.(\varphi \Rightarrow \psi)$. □

We have presented the semantics for intuitionistic predicate calculus in a form which largely follows that of [Rasiowa and Sikorski]. Almost all the results mentioned in this chapter can be found in chapters eight to ten of that book. This approach allows a presentation which is quite uniform with that given for the classical and Boolean valued cases above. However, there are a number of different (though of course actually equivalent) semantics for first order intuitionistic logic. A good introduction to intuitionistic logic via Kripke semantics, which is probably the most commonly considered semantics nowadays, can be found in [Nerode]. Some investigators concern themselves with the effect of restricting metalinguistic reasoning to intuitionistically acceptable patterns. An interesting discussion can be found in [Dummett].

CHAPTER XI

THE INTUITIONISTIC CASE: ε AND τ

Readers who are still with us will no doubt by now have noted a peculiarity. Once we left the λ-calculus behind we presented a syntax for arbitrary term forming operators, then proceeded to ignore operators other than ε for several chapters. Part of the reason for this is that other familiar operators don't fit well with our concerns: ι is mathematically uninteresting; set abstraction is defined only for particular theories; η is only well suited to formulation in free logic. The best candidate for our consideration is τ, and of course we had a perfectly good reason for ignoring τ, too. For in classical logic we can define $\tau x.\varphi \equiv \varepsilon x.\neg\varphi$, and so we have implicitly already covered the classical τ-calculi.

Now that we are moving on to deal with term forming operators in intuitionistic logic this definition will no longer do, as we shall see. We will therefore have to deal explicitly and separately with the ε- and τ-calculi in this chapter. Fortunately we already have in place many of the elements we will need to give our basic semantic definitions and prove our basic semantic theorems more briefly and efficiently than we have been able to in earlier chapters.

What makes the case of term forming operators, particularly ε, in intuitionistic logic interesting is that the ε-theorems do not hold—which is to say that the introduction of these operators is non-conservative. And the new consequences that follow from their addition are interesting, including as they do schemes (such as deMorgan's law) which are commonly thought by those in the thrall of classical logic to be logical principles.

We will proceed as follows in this chapter. First we will set out the various systems we will work with. We will then go straight on to introducing the semantics and proving the Soundness and Completeness Theorems for these systems. The semantics we will introduce will play a role similar to that played by the 'syntactical' semantics in our treatment of the λ-calculus. That is, these structures are more or less the structures which are required by the ε- and τ-axioms. At least, this is the reasonable way to look at them if we regard the extensionality assumptions as a device designed to allow us to simplify our semantics. We can then go on to use these 'syntactical' structures as a tool for investigating the more interesting and more mathematical features of these structures, and as a tool for investigating the question of what can and cannot be proved in these calculi, in particular for constructing independence proofs. And, of course, we can hope to make these 'syntactical' structures slicker as we learn more about them, much as happened in the λ case—though, admittedly, nowhere near the same kind of progress has yet been made in the current case.

In the scant literature on the ε-operator in intuitionistic logic, most researchers have begun by noting that the ε-theorems do not hold in intuitionistic logic. From there some authors have set out to discover what sorts of modifications will suffice to make them hold. Sometimes these modifications are restrictions on the class of sentences to which the theorem is applied, e. g. [Rasiowa 1956]. Others instead restrict the notion of derivation in some way, so that the offending formulas can no longer be proved, e. g. [Mints, 1974]. Most of the work in this literature involves the use of proof theoretic methods (e. g. [Mints 1974, 1990, (to appear)], and apparently also work by Maehara and Shirai in Japan and Smirnov in the former USSR, see [Mints 1990] for a discussion of their work). However, [Rasiowa 1956], employs model theoretic methods closely related to those used here.

The other strand in this literature seems to consist pretty much exclusively of [Bell 1993a, 1993b]. For the first order case Bell supplies a semantics and gives a soundness proof so that he can prove some independence results for the intuitionistic ε–calculus. However, he does not provide a completeness proof for the very good reason that none of the systems he considers is complete for the semantics he uses.

The completeness proofs we give in this chapter are the first in the literature for the intuitionistic ε–calculus. While they are perhaps not so elegant they will allow us to present most all the other results known about the first order intuitionistic calculus as an interrelated package, as well as allowing us to prove some new independence results. This is also the first extended treatment of the intuitionistic τ–calculus.

1. First Order Intuitionistic ε–Calculi and τ–Calculi

First, some terminology. Since we are now dealing with Heyting algebras and not only Boolean algebras, we will rename (BAck) from Chapter 9:

$$\text{(HAck)} \qquad \forall x.[(\exists y.\varphi \Rightarrow \varphi[y/x]) \Leftrightarrow (\exists y.\psi \Rightarrow \psi[y/x])] \Rightarrow \varepsilon y.\varphi = \varepsilon y.\psi.$$

(This axiom scheme is not to be confused with the present author.) Likewise we will rename (B$*$) as

$$(\varepsilon\text{H}^*) \qquad \frac{\forall x.[(\exists y.\varphi \Rightarrow \varphi[y/x]) \Leftrightarrow (\exists y.\psi \Rightarrow \psi[y/x])]}{\varepsilon y.\varphi = \varepsilon y.\psi}.$$

We need also the corresponding principles for the τ–calculus. So we introduce

$$(\tau\text{Hext}) \qquad \forall x.[(\varphi[y/x] \Rightarrow \forall y.\varphi) \Leftrightarrow (\psi[y/x] \Rightarrow \forall y.\psi)] \Rightarrow \tau y.\varphi = \tau y.\psi.$$

and

$$(\tau\text{H}^*) \qquad \frac{\forall x.[(\varphi[y/x] \Rightarrow \forall y.\varphi) \Leftrightarrow (\psi[y/x] \Rightarrow \forall y.\psi)]}{\tau y.\varphi = \tau y.\psi}.$$

Definition 1. The *H-extensional intuitionistic ε-calculus* is the system which results if we replace (Ack) by (HAck) in the extensional intuitionistic ε–calculus. Similarly, the *H-extensional intuitionistic τ-calculus* results if we replace (τext) by (τHext) in the extensional intuitionistic τ–calculus.

The *intuitionistic Hε^*-calculus* is the calculus which results if we replace the rule (ε^*) by (Hε^*) in the intuitionistic ε^*–calculus. Similarly, the *intuitionistic Hτ^*-calculus* is the calculus which results if we replace (τ^*) by (Hτ^*) in the intuitionistic τ^*–calculus.

We will be primarily concerned with six different calculi here: the H–extensional and H* versions of both the intuitionistic ε and τ calculi, and the non-extensional versions of each. However, we will occasionally have cause to mention the ε^*–calculus (τ^*–calculus) and the extensional ε–calculus (τ–calculus). So we will need a variety of different notations for the claim that φ is derivable from Σ in these various calculi. So we will write '$\Sigma \vdash_{H\varepsilon^*} \varphi$' for that claim in the intuitionistic Hε^*–calculus, '$\Sigma \vdash_{H\varepsilon ext} \varphi$' for the H–extensional intuitionistic ε–calculus case, '$\Sigma \vdash_{\varepsilon} \varphi$' for the intuitionistic ε–calculus, and '$\Sigma \vdash_{H\tau^*} \varphi$', '$\Sigma \vdash_{H\tau ext} \varphi$', and '$\Sigma \vdash_{\tau} \varphi$' for the corresponding τ cases. On those occasions where we need to discuss them, we will write '$\Sigma \cup (Ack) \vdash_{\varepsilon} \varphi$' and '$\Sigma \cup (\tau ext) \vdash_{\tau} \varphi$' for claims of derivability in the extensional intuitionistic calculi, and '$\Sigma \vdash_{\varepsilon^*} \varphi$' and $\Sigma \vdash_{\tau^*} \varphi$ for such claims in the intuitionistic ε^*– and τ^*–calculi.

Since we will be dealing with several different calculi we will need to consider several different kinds of structures. First we will introduce the structures we need for the cases where we have some sort of extensionality condition.

Definition 2. We recall that in Definition 9.2 we said that a join of form $\bigvee_{d \in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ is *attained in* a Boolean valued structure $\mathcal{M}$ if there is a $d \in D$ such that $\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \bigvee_{d \in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$. We extend this definition to Heyting algebra valued structures, and we also say that *meets* of form $\bigwedge_{d \in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ are *attained in* $\mathcal{M}$ if $\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \bigwedge_{d \in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ for some $d \in D$.

We will say that a structure $\mathcal{M}$ for $\mathcal{L}'$ is *n–join–attained* (*n–meet–attained*) for $n \in \omega$ if for any formula φ containing n or fewer occurrences of ε (of τ) and all valuations ϱ, the join $\bigvee_{d \in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ (the meet $\bigwedge_{d \in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$) is attained.

Let E be a choice function on $\mathcal{P}(D)$, and $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c}\rangle$ be an intuitionistic $\mathcal{L}'$ structure.

(A) In suitable $\mathcal{M}$, we can extend the definitions of $[t]^{\varrho}_{\mathcal{M}}$ and $\|\varphi\|^{\varrho}_{\mathcal{M}}$ from the terms and formulas of $\mathcal{L}'$ to the whole of **Wffs** and **Terms** in the ε language $\mathcal{L}$ by rewriting Definition 10.1, except that we write $\mathcal{L}$ instead of $\mathcal{L}'$ and add the clause that if $\mathcal{M}$ is n–join–attained then for φ with n or fewer occurrences of ε,

$$\|\varepsilon x.\varphi\|^{\varrho}_{\mathcal{M}} = E\left(\{d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}}\}\right).$$

If $\mathcal{M}$ is n–join–attained for all $n \in \omega$, we say that $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c}, E\rangle$ is a *quasi-extensional intuitionistic ε-structure* for $\mathcal{L}$.

(B) In suitable $\mathcal{M}$, we can extend the definitions of $[t]^{\varrho}_{\mathcal{M}}$ and $\|\varphi\|^{\varrho}_{\mathcal{M}}$ from the terms and formulas of $\mathcal{L}'$ to the whole of **Wffs** and **Terms** in the τ language $\mathcal{L}$ by rewriting Definition 10.1, except that we write $\mathcal{L}$ instead of $\mathcal{L}'$ and add the clause that if $\mathcal{M}$ is n–meet–attained then for φ with n or fewer occurrences of τ,

$$\|\tau x.\varphi\|^{\varrho}_{\mathcal{M}} = E\left(\{d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigwedge_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}}\}\right).$$

We say that $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c}, E\rangle$ is a *quasi-extensional intuitionistic τ-structure* for $\mathcal{L}$ if $\mathcal{M}$ is n–meet–attained for all $n \in \omega$.

Now the following theorems are easily proved using the same proofs as were used for Theorems 7.1 and 7.2, except that in a few places we must refer to $\{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} \,\}$ (or $\{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigwedge_{d \in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} \,\}$) rather than simply $\{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = 1 \,\}$.

Theorem 1. *Let $\mathcal{M}$ be a quasi-extensional ε-structure (τ-structure) for $\mathcal{L}$, M a well-formed expression, and ϱ, ϱ' two valuations in $\mathcal{M}$ such that for $x \in \mathbf{FV}(M)$, $\varrho(x) = \varrho'(x)$. Then*

(1) $M \equiv t \in \mathbf{Terms} \implies [t]^{\varrho}_{\mathcal{M}} = [t]^{\varrho'}_{\mathcal{M}}$.

(2) $M \equiv \varphi \in \mathbf{Wffs} \implies \|\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho'}_{\mathcal{M}}$. □

Theorem 2. *Let M be a well-formed expression of $\mathcal{L}$, $\mathcal{M}$ a quasi-extensional intuitionistic ε-structure (τ-structure) for $\mathcal{L}$, ϱ a valuation in $\mathcal{M}$, and $s \in$ **Terms**. Then*

(1) $M \equiv t \in \mathbf{Terms} \implies [t]^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})}_{\mathcal{M}} = [t[x/s]]^{\varrho}_{\mathcal{M}}$.

(2) $M \equiv \varphi \in \mathbf{Wffs} \implies \|\varphi\|^{\varrho(x/[s]^{\varrho}_{\mathcal{M}})}_{\mathcal{M}} = \|\varphi[x/s]\|^{\varrho}_{\mathcal{M}}$. □

Let $\mathcal{M}$ be a quasi-extensional intuitionistic ε– (or τ–) structure for $\mathcal{L}$. A formula φ of $\mathcal{L}$ is *satisfiable in* $\mathcal{M}$ if $\|\varphi\|^{\varrho}_{\mathcal{M}} = 1$ for some valuation ϱ in $\mathcal{M}$. We write '$\mathcal{M}, \varrho \models_q \varphi$' for this. We say that φ is *valid in* $\mathcal{M}$ and write '$\mathcal{M} \models_q \varphi$' if $\mathcal{M}, \varrho \models_q \varphi$ for all ϱ. If $\mathcal{M} \models_q \varphi$ for all $\mathcal{M}$, we say that φ is valid, and write '$\models_q \varphi$'. If $\Sigma \subseteq$ **Wffs**, we will write $\mathcal{M} \models_q \Sigma$ if $\mathcal{M} \models_q \beta$ for each $\beta \in \Sigma$. We write $\Sigma \models_q \varphi$ if

$$\mathcal{M} \models_q \Sigma \implies \mathcal{M} \models_q \varphi.$$

Just as in Chapter 9, we will need to distinguish between *extensional* and *quasi-extensional* structures. We will say that an *extensional intuitionistic ε-structure* is a punctilious quasi-extensional intuitionistic ε-structure, i.e. if $\mathcal{M} \models_q (HAck)$, we will call $\mathcal{M}$ an extensional intuitionistic ε–structure. We will write '$\models$' instead of '$\models_q$' if we are restricting attention to extensional intuitionistic ε–structures.

We will also need the corresponding notion for τ. So we will say that a quasi-extensional intuitionistic τ–structure is τ–punctilious if for all φ, ψ, and ϱ,

$$\bigwedge_{d\in D} \|(\varphi[y/x] \Rightarrow \forall y.\varphi) \Leftrightarrow (\psi[y/x] \Rightarrow \forall y.\psi)\|_{\mathcal{M}}^{\varrho(x/d)} \leqslant \mathrm{eq}_{\mathcal{M}}([\tau y.\varphi]_{\mathcal{M}}^{\varrho}, [\tau y.\psi]_{\mathcal{M}}^{\varrho}),$$

and we define an *extensional intuitionistic τ-structure* to be a τ–punctilious quasi-extensional intuitionistic τ–structure. Once again, we will write '$\models$' instead of '$\models_q$' when we are restricting attention to this case.

Theorem 3 (Soundness Theorems). (1) $\Sigma \vdash_{\varepsilon^*} \varphi \implies \Sigma \models_q \varphi$.

(2) $\Sigma \vdash_{H\varepsilon^*} \varphi \implies \Sigma \models_q \varphi$.

(3) $\Sigma \cup (Ack) \vdash_{\varepsilon} \varphi \implies \Sigma \models \varphi$.

(4) $\Sigma \vdash_{H\varepsilon ext} \varphi \implies \Sigma \models \varphi$.

(1′) $\Sigma \vdash_{\tau^*} \varphi \implies \Sigma \models_q \varphi$.

(2′) $\Sigma \vdash_{H\tau^*} \varphi \implies \Sigma \models_q \varphi$.

(3′) $\Sigma \cup (\tau ext) \vdash_{\tau} \varphi \implies \Sigma \models \varphi$.

(4′) $\Sigma \vdash_{H\tau ext} \varphi \implies \Sigma \models \varphi$.

Proof. By Theorem 10.1 and the fact that a quasi-extensional intuitionistic ε– or τ–structure is an intuitionistic structure, it suffices to show that the rules and axioms newly introduced in each system are valid. We have noted that in all our systems the analogue of Lemma 7.4 holds, so (α) is valid in all these systems. Our definitions ensure in the various ε–cases that

$$\|\exists x.\varphi\|_{\mathcal{M}}^{\varrho} = \|\varphi[x/\varepsilon x.\varphi]\|_{\mathcal{M}}^{\varrho}$$

and in the τ–cases that

$$\|\forall x.\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi[x/\tau x.\varphi]\|^{\varrho}_{\mathcal{M}}\,,$$

and it follows that the (ε) axioms are valid in the various ε–structures and the (τ) axioms are valid in the various τ–structures.

For (4) and (4′), the validity of (HAck) and (Hτext) is obvious by the definition of extensional structures. (3) and (3′) follow since (HAck)$\vdash_{IC}$(Ack) and Hτext$\vdash_{IC}$(τext).

Now let

$$A = \{\, d \in D \mid \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d\in D} \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} \,\}$$

and

$$B = \{\, d \in D \mid \|\psi\|^{\varrho(x/d)}_{\mathcal{M}} = \bigvee_{d\in D} \|\psi\|^{\varrho(x/d)}_{\mathcal{M}} \,\}.$$

Define A' and B' similarly, replacing $\bigvee$ with $\bigwedge$. But if $\|\forall x.(\varphi \Leftrightarrow \psi)\|^{\varrho}_{\mathcal{M}} = 1$, $A = B$ and $A' = B'$ follow, and since E is a function $A = B \implies \|\varepsilon x.\varphi = \varepsilon x.\psi\|^{\varrho}_{\mathcal{M}} = 1$ and $A' = B' \implies \|\tau x.\varphi = \tau x.\psi\|^{\varrho}_{\mathcal{M}} = 1$, so (1) and (1′) are proved. Finally, $A = B$ if $\|\forall x.[(\exists y.\varphi \Rightarrow \varphi[y/x]) \Leftrightarrow (\exists y.\psi \Rightarrow \psi[y/x])]\|^{\varrho}_{\mathcal{M}} = 1$, and $A' = B'$ if $\|\forall x.[(\varphi[y/x] \Rightarrow \forall y.\varphi) \Leftrightarrow (\psi[y/x] \Rightarrow \forall y.\psi)]\|^{\varrho}_{\mathcal{M}} = 1$, and so (2) and (2′) follow. □

Now in order to prove our completeness results we will again use the Lindenbaum algebra construction. We will write '$\vdash$' to stand ambiguously for $\vdash_{H\varepsilon^\bullet}$, $\vdash_{H\tau^\bullet}$, $\vdash_{H\varepsilon ext}$ and $\vdash_{H\tau ext}$ for the parts that are common to all of the proofs. As usual we put $\varphi \approx \psi \iff \Sigma \vdash \varphi \Leftrightarrow \psi$, and $\|\varphi\| = \{\, \psi \in \mathbf{Wffs} \mid \varphi \approx \psi \,\}$. We put $\Omega = \{\|\varphi\| \mid \varphi \in \mathbf{Wffs}\,\}$ and order it by $\|\varphi\| \leqslant \|\psi\| \iff \Sigma \vdash \varphi \Rightarrow \psi$. The usual proof shows that Ω

is a Heyting algebra, and we say that Ω is the Lindenbaum algebra for Σ. As usual, we can prove that $\|\exists x.\varphi\| = \bigvee_{t\in\mathbf{Terms}} \|\varphi[x/t]\|$ and $\|\forall x.\varphi\| = \bigwedge_{t\in\mathbf{Terms}} \|\varphi[x/t]\|$.

Now we put, for s, $t \in \mathbf{Terms}$, $s \sim t \iff \Sigma \vdash s = t$, and put $[t] = \{ s \mid s \sim t \}$. We will define $[\mathbf{Terms}]$ to be $\{ [t] \mid t \in \mathbf{Terms} \}$. We can define, for $P_i \in \mathbf{Pred}_n$,

$$\mathcal{R}_i(\langle [t_1], \ldots, [t_n] \rangle) = \|P_i t_1 \ldots t_n\|,$$

and

$$\mathrm{eq}_{\mathcal{M}}([t_1], [t_2]) = \|t_1 = t_2\|$$

and show in the routine way that this is independent of choice of representative for the $[t_i]$. Similarly, we can put for $j \in J$

$$\mathbf{c}(j) = [c_j].$$

It is straightforward to show that in the Lindenbaum algebra

$$\bigvee_{t\in\mathbf{Terms}} \|\varphi[x/t]\| = \bigvee_{[t]\in[\mathbf{Terms}]} \|\varphi[x/t]\|$$

and

$$\bigwedge_{t\in\mathbf{Terms}} \|\varphi[x/t]\| = \bigwedge_{[t]\in[\mathbf{Terms}]} \|\varphi[x/t]\|.$$

What remains to be defined for us to be able to have an intuitionistic ε–or τ–structure is a choice function E.

Definition 3. Let $\varphi \in \mathbf{Wffs}$ with free variables among $\{x, z_1, \ldots, z_n\}$, and $t_1, \ldots, t_n \in \mathbf{Terms}$. We use the abbreviations $\vec{z} = \langle z_1, \ldots, z_n \rangle$ and $\vec{t} = \langle t_1, \ldots, t_n \rangle$.

Let $N \subseteq [\mathbf{Terms}]$. We say that the triple $\langle \varphi, x, \vec{t} \rangle$ *ε-represents* N (and that N is *ε-represented* by the triple, and so that N is *ε-representable*) if

$$N = \{ [s] \in [\mathbf{Terms}] \mid \|\varphi[x/s, \vec{z}/\vec{t}]\| = \bigvee_{[s] \in [\mathbf{Terms}]} \|\varphi[x/s, \vec{z}/\vec{t}]\| \}.$$

We say that the triple $\langle \varphi, x, \vec{t} \rangle$ *τ-represents* N (and that N is *τ-represented* by the triple, and so that N is *τ-representable*) if

$$N = \{ [s] \in [\mathbf{Terms}] \mid \|\varphi[x/s, \vec{z}/\vec{t}]\| = \bigwedge_{[s] \in [\mathbf{Terms}]} \|\varphi[x/s, \vec{z}/\vec{t}]\| \}.$$

We need

Lemma 1. (1) *If $\langle \varphi, x, \vec{t} \rangle$ and $\langle \psi, x', \vec{t'} \rangle$ both ε-represent $N \subseteq [\mathbf{Terms}]$, then $\Sigma \vdash \varepsilon x.\varphi = \varepsilon x'.\psi$.*

(2) *If $\langle \varphi, x, \vec{t} \rangle$ and $\langle \psi, x', \vec{t'} \rangle$ both τ-represent $N \subseteq [\mathbf{Terms}]$, then $\Sigma \vdash \tau x.\varphi = \tau x'.\psi$.*

Proof. Unfortunately we cannot simply duplicate the proof of the similar Lemma 9.4 to prove (1) because that proof appealed to Lemma 7.11(b), which is true for classical ε–calculi but is not in general true for intuitionistic ε–calculi. Fortunately this is just an inconvenience, not a disaster.

We have $\mathbf{FV}(\varphi) \subseteq \{x, z_1, \ldots, z_n\}$ and $\mathbf{FV}(\psi) \subseteq \{x', z'_1, \ldots, z'_l\}$ for some n and l and we can assume without loss of generality that all these variables are distinct. We can come up with a sequence of $n + l + 1$ variables which occur nowhere in φ or ψ, $\vec{w} = \langle w, w_1, \ldots, w_{l+n} \rangle$ and a sequence of terms $\vec{s} = \langle t_1, \ldots, t_n, t'_1, \ldots, t'_l \rangle$. Now if we take $\varphi' \equiv \varphi[x/w, z_1/w_1, \ldots, z_n/w_n]$ and $\psi' = \psi[x/w, z'_1/w_{n+1}, \ldots, z'_l/w_{n+l}]$ it is easily seen that for any $t \in \mathbf{Terms}$

$$\varphi'[w/t, \vec{w}/\vec{s}] \equiv \varphi[x/t, \vec{z}/\vec{t}]$$

and

$$\psi'[w/t, \vec{w}/\vec{s}] \equiv \psi[x/t, \vec{z'}/\vec{t'}].$$

So N is represented by both $\langle \varphi', w, \vec{s} \rangle$ and $\langle \psi', w, \vec{s} \rangle$.

Now for an arbitrary $r \in \mathbf{Terms}$

$$r \in N \iff \|\varphi'[w/r, \vec{w}/\vec{s}]\| = \bigvee_{[s]\in[\mathbf{Terms}]} \|\varphi'[w/s, \vec{w}/\vec{s}]\|$$

and

$$r \in N \iff \|\psi'[w/r, \vec{w}/\vec{s}]\| = \bigvee_{[s]\in[\mathbf{Terms}]} \|\psi'[w/s, \vec{w}/\vec{s}]\|,$$

which is to say $\Sigma \vdash \varphi'[w/r, \vec{w}/\vec{s}] \Leftrightarrow \exists w.\varphi'[\vec{w}/\vec{s}]$ if and only if $\Sigma \vdash \psi'[w/r, \vec{w}/\vec{s}] \Leftrightarrow \exists w.\psi'[\vec{w}/\vec{s}]$. It follows by the deduction theorem (twice) and the definition of $\Leftrightarrow$ that

$$\Sigma \vdash (\varphi'[w/r, \vec{w}/\vec{s}] \Leftrightarrow \exists w.\varphi'[\vec{w}/\vec{s}]) \Leftrightarrow (\psi'[w/r, \vec{w}/\vec{s}] \Leftrightarrow \exists w.\psi'[\vec{w}/\vec{s}]),$$

whence

$$\Sigma \vdash (\exists w.\varphi'[\vec{w}/\vec{s}] \Rightarrow \varphi'[w/r, \vec{w}/\vec{s}]) \Leftrightarrow (\exists w.\psi'[\vec{w}/\vec{s}] \Rightarrow \psi'[w/r, \vec{w}/\vec{s}]).$$

Since this holds for arbitrary r, it will also hold for a fresh variable v (there is such a fresh variable since Σ is a set of sentences and since $\mathbf{FV}(\varphi)$ and $\mathbf{FV}(\psi)$ are finite), and so we have by $\forall$-introduction that

$$\Sigma \vdash \forall x.[(\exists w.\varphi'[\vec{w}/\vec{s}] \Rightarrow \varphi'[\vec{w}/\vec{s}]) \Leftrightarrow (\exists w.\psi'[\vec{w}/\vec{s}] \Rightarrow \psi'[\vec{w}/\vec{s}])],$$

and so by (*) or (BAck), as the case may be, we have

$$\varepsilon x.\varphi[\vec{z}/\vec{t}] \equiv \varepsilon x.\varphi'[\vec{w}/\vec{s}] = \varepsilon x.\psi'[\vec{w}/\vec{s}] = \varepsilon x'\psi[\vec{w}/\vec{s}]$$

with the last equality justified by (α).

The modifications that need to be made to adjust this proof for (2) are obvious. □

So we can define our choice function $E : \mathcal{P}([\mathbf{Terms}]) \to [\mathbf{Terms}]$ by putting

$$E(N) = [\varepsilon x.\varphi[\vec{z}/\vec{t}]]$$

$$(E(N) = [\varepsilon x.\varphi[\vec{z}/\vec{t}]])$$

if N is ($\varepsilon-$ or $\tau-$) represented by $\langle\varphi, x, \vec{t}\rangle$, and if $N \neq \emptyset$ is not representable, we can let $E(N)$ be an arbitrary member of N. The (ε) or (τ) axioms ensure that this is a choice function, as usual. The reader will notice that we have defined our current structures so that E need not assign anything to $\emptyset$ (which is obviously okay since we require the joins (meets) used to interpret existential (universal) quantifiers to be attained), so this definition is sufficient.

We have now almost defined the structure we want. Again, the problem is that we have no guarantee that Ω is complete. So again we appeal to Proposition 10.1. Let $i : \Omega \to \Omega'$ be a selected injection of the type guaranteed to exist by that proposition. $\mathcal{M} = \langle[\mathbf{Terms}], \Omega', \{\, i \circ \mathcal{R}_i : i \in I\}, i \circ \mathrm{eq}_{\mathcal{M}}, \mathbf{c}, E\rangle$ is easily seen to be a quasi-extensional intuitionistic ($\varepsilon-$ or $\tau-$) structure for $\mathcal{L}$, and if we are discussing an H–extensional calculus it will obviously be an extensional structure. If we consider the map $\pi : \mathbf{Var} \to [\mathbf{Terms}]$ defined by $\pi(x) = [x]$, it is a straightforward matter to show that $\Sigma \vdash \varphi \iff \|\varphi\|^{\pi}_{\mathcal{M}} = 1$. We call π the canonical valuation.

Theorem 4 (Completeness Theorems). *In each of the following groups of three claims the claims are mutually equivalent.*

(1) $\Sigma \vdash_{H\varepsilon^*} \varphi$.

(2) $\|\varphi\|^{\pi}_{\mathcal{M}} = 1$ *on the canonical valuation.*

(3) $\Sigma \models_q \varphi$.

(1′) $\Sigma \vdash_{H\varepsilon ext} \varphi$.

(2′) $\|\varphi\|^{\pi}_{\mathcal{M}} = 1$ *on the canonical valuation.*

(3′) $\Sigma \models \varphi$.

(1″) $\Sigma \vdash_{H\tau^*} \varphi$.

(2″) $\|\varphi\|^{\pi}_{\mathcal{M}} = 1$ *on the canonical valuation.*

(3″) $\Sigma \models_q \varphi$.

(1‴) $\Sigma \vdash_{H\tau ext} \varphi$.

(2‴) $\|\varphi\|^{\pi}_{\mathcal{M}} = 1$ *on the canonical valuation.*

(3‴) $\Sigma \models \varphi$.

Proof. The proof is the same in all four cases. (3) $\Longrightarrow$ (2) is by definition of $\models$ and the fact that $\mathcal{M}$ is a structure. (1) $\Longleftrightarrow$ (2) is by construction, as just mentioned, and (1) $\Longrightarrow$ (3) is Soundness. □

It remains to give a semantics for the intuitionistic ε– and τ–calculi which eliminates the need to add extensionality assumptions of any sort. The key to doing so is, not surprisingly, to be found in §2 of Chapter 8.

To begin with, the reader is reminded of the definition, for each ε–term t, of the *skeleton term* $\kappa(t)$ for t (Definition 8.7). We will use the same definition, writing τ for ε, to define the skeleton term of each τ-term, and will also use '$\kappa(t)$' to denote the skeleton of any τ–term t.

We also remind the reader of the definition of an ε*-choice function* (Definition 8.8). We will now sometimes call an ε–choice function a τ*-choice function*.

Definition 4. Let $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c}\rangle$ be an intuitionistic structure for $\mathcal{L}'$ and E an ε–choice function (a τ–choice function). A join of form $\bigvee_{d\in D}\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ is *attained* if there is a $d \in D$ such that $\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \bigvee_{d\in D}\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$. (A meet of form $\bigwedge_{d\in D}\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ is *attained* if there is a $d \in D$ such that $\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \bigwedge_{d\in D}\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$).

A structure is n–join attained (n–meet attained) for $n \in \omega$ if for any φ with n or fewer occurrences of ε (of τ) and all ϱ the join $\bigvee_{d\in D}\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ is attained (the meet $\bigwedge_{d\in D}\|\varphi\|_{\mathcal{M}}^{\varrho(x/d)}$ is attained).

For every φ with n or fewer occurrences of ε, we can extract an ordering of the free *occurrences* of variables in $\varepsilon x.\varphi$ from the procedure used to find its skeleton $\kappa(\varepsilon x.\varphi)$. Write $\langle y_1, \ldots, y_n\rangle$ for the free occurrences of variables in $\varepsilon x.\varphi$, and $\langle x_1, \ldots, x_n\rangle$ for the free occurrences of variables in $\kappa(\varepsilon x.\varphi)$. If $\kappa(\varepsilon x.\varphi) \equiv \varepsilon v_1.\beta$, let M be the set:

$$\{\, d \in D \mid \|\beta\|_{\mathcal{M}}^{\varrho(v_1/d)(x_1/\varrho(y_1))\ldots(x_n/\varrho(y_n))} = \bigvee_{d\in D}\|\beta\|_{\mathcal{M}}^{\varrho(v_1/d)(x_1/\varrho(y_1))\ldots(x_n/\varrho(y_n))}\,\}.$$

If $\mathcal{M}$ is n–join–attained, we know that $M \neq \emptyset$.

Now we can extend the definitions of $[t]_{\mathcal{M}}^{\varrho} \in D$ and $\|\varphi\|_{\mathcal{M}}^{\varrho} \in \Omega$ from the terms and formulas of $\mathcal{L}'$ to all of **Terms** and **Wffs** by rewriting Definition 10.1, except that we write $\mathcal{L}$ for $\mathcal{L}'$, and we add the following clause. If $\mathcal{M}$ is n-join-attained, then for any φ with n or fewer occurrences of ε and for any x

$$[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} = E(M, \kappa(\varepsilon x.\varphi), \langle \varrho(y_1), \ldots, \varrho(y_n)\rangle).$$

(For every φ with n or fewer occurrences of τ, we can extract an ordering of the free *occurrences* of variables in $\tau x.\varphi$ from the procedure used to find its skeleton $\kappa(\tau x.\varphi)$. Write $\langle y_1, \ldots, y_n\rangle$ for the free occurrences of variables in $\tau x.\varphi$, and $\langle x_1, \ldots, x_n\rangle$ for the free occurrences of variables in $\kappa(\tau x.\varphi)$. If $\kappa(\tau x.\varphi) \equiv \varepsilon v_1.\beta$, let M' be the set:

$$\{\, d \in D \mid \|\beta\|^{\varrho(v_1/d)(x_1/\varrho(y_1))\ldots(x_n/\varrho(y_n))}_{\mathcal{M}} = \bigwedge_{d\in D} \|\beta\|^{\varrho(v_1/d)(x_1/\varrho(y_1))\ldots(x_n/\varrho(y_n))}_{\mathcal{M}} \,\}.$$

If $\mathcal{M}$ is n-meet-attained, we know that $M' \neq \emptyset$.

Now we can extend the definitions of $[t]^{\varrho}_{\mathcal{M}} \in D$ and $\|\varphi\|^{\varrho}_{\mathcal{M}} \in \Omega$ from the terms and formulas of $\mathcal{L}'$ to all of **Terms** and **Wffs** by rewriting Definition 10.1, except that we write $\mathcal{L}$ for $\mathcal{L}'$, and we add the following clause. If $\mathcal{M}$ is n-meet-attained, then for any φ with n or fewer occurrences of τ and for any x

$$[\tau x.\varphi]^{\varrho}_{\mathcal{M}} = E(M', \kappa(\tau x.\varphi), \langle \varrho(y_1), \ldots, \varrho(y_n)\rangle).)$$

If $\mathcal{M}$ is n-join-attained for all $n \in \omega$, $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathrm{c}, E\rangle$ is an *intuitionistic ε-structure*. If $\mathcal{M}$ is n-meet-attained for all $n \in \omega$, $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathrm{c}, E\rangle$ is an *intuitionistic τ-structure*.)

It is a straightforward matter to prove an analogue of Theorem 1 for intuitionistic ε- and τ- structures. However, as was noted in Chapter 8 we cannot get a precise analogue to Theorem 2 if we give up our extensionality assumptions. Luckily, the proof of Theorem 8.12 does not depend on the fact that we are working in classical rather than intuitionistic logic, so we can get

Theorem 5. *Let M be a well-formed expression of $\mathcal{L}$, $\mathcal{M}$ an intuitionistic ε–structure (an intuitionistic τ–structure) for $\mathcal{L}$, ϱ a valuation in $\mathcal{M}$, and $s \in$* **Terms**. *Then*

(1) $M \equiv t \in \textbf{Terms} \implies [t]_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = [t[x/s]]_{\mathcal{M}}^{\varrho}$

(2) $M \equiv \varphi \in \textbf{Wffs} \implies \|\varphi\|_{\mathcal{M}}^{\varrho(x/[s]_{\mathcal{M}}^{\varrho})} = \|\varphi[x/s]\|_{\mathcal{M}}^{\varrho}$

provided that x does not have a free occurrence in M which is a substring of an ε–term (τ–term) which is a proper subterm of M. □

If $\mathcal{M}$ is an intuitionistic ε–structure, we write '$\mathcal{M}, \varrho \models_\varepsilon \varphi$' if $\|\varphi\|_{\mathcal{M}}^{\varrho} = 1$ and '$\mathcal{M} \models_\varepsilon \varphi$' if $\mathcal{M}, \varrho \models_\varepsilon \varphi$ for all ϱ. We write '$\mathcal{M} \models_\varepsilon \Sigma$' if $\mathcal{M} \models_\varepsilon \beta$ for all $\beta \in \Sigma$. We write '$\Sigma \models_\varepsilon \varphi$' if $\mathcal{M} \models_\varepsilon \Sigma$ implies $\mathcal{M} \models_\varepsilon \varphi$. If $\mathcal{M}$ is an intuitionistic τ–structure, we write '$\mathcal{M}, \varrho \models_\tau \varphi$' if $\|\varphi\|_{\mathcal{M}}^{\varrho} = 1$ and '$\mathcal{M} \models_\tau \varphi$' if $\mathcal{M}, \varrho \models_\tau \varphi$ for all ϱ. We write '$\mathcal{M} \models_\tau \Sigma$' if $\mathcal{M} \models_\tau \beta$ for all $\beta \in \Sigma$. We write '$\Sigma \models_\tau \varphi$' if $\mathcal{M} \models_\tau \Sigma$ implies $\mathcal{M} \models_\tau \varphi$.

Theorem 6 (Soundness Theorem). (1) $\Sigma \vdash_\varepsilon \varphi \implies \Sigma \models_\varepsilon \varphi$.

(2) $\Sigma \vdash_\tau \varphi \implies \Sigma \models_\tau \varphi$.

Proof. As usual we only need to prove that (ε) and (α) ((τ) and (α)) are valid. But the relevant substitution lemma holds, as we have already mentioned, so (α) is valid. And our definitions ensure that $\|\varphi[x/\varepsilon x.\varphi]\|_{\mathcal{M}}^{\varrho} = \|\exists x.\varphi\|_{\mathcal{M}}^{\varrho}$ and $\|\varphi[x/\tau x.\varphi]\|_{\mathcal{M}}^{\varrho} = \|\forall x.\varphi\|_{\mathcal{M}}^{\varrho}$. □

The easiest way to prove our completeness result now is to once again carry out the steps of the canonical model proof outlined above, up until the time comes to define our choice function E, which now needs to be an ε–choice function (or a τ–choice function). We can leave the definition of representable sets unchanged. If $N \subseteq [\mathbf{Terms}]$ is represented by $\langle \varphi, y, \vec{t} \rangle$, and if $\langle x_1, \ldots, x_n \rangle$ is the sequence of free occurrences of variables in $(\varepsilon y.\varphi)[\vec{z}/\vec{t}]$, then we can define E by

$$E\left(N, \kappa\left((\varepsilon y.\varphi)[\vec{z}/\vec{t}]\right), \langle [x_1], \ldots, [x_n] \rangle\right) = [(\varepsilon y.\varphi)[\vec{z}/\vec{t}]].$$

(if $\langle x_1, \ldots, x_n \rangle$ is the sequence of free occurrences of variables in $(\tau y.\varphi)[\vec{z}/\vec{t}]$, then we can define E by

$$E\left(N, \kappa\left((\tau y.\varphi)[\vec{z}/\vec{t}]\right), \langle [x_1], \ldots, [x_n] \rangle\right) = [(\tau y.\varphi)[\vec{z}/\vec{t}]]).$$

For unrepresentable N, or ground terms and/or sequences that do not represent N, we can let E choose an arbitrary element of $[\mathbf{Terms}]$. To show that this is well-defined we need to show, writing $\vec{y} = \langle y_1, \ldots, y_n \rangle$ for the sequence of free occurrences of variables in $(\varepsilon x.\psi)[\vec{z}/\vec{d}]$, the following: if

(1) the set N is represented by both $\langle \varphi, y, \vec{s} \rangle$ and $\langle \psi, x, \vec{t} \rangle$;

(2) $\kappa((\varepsilon y.\varphi)[\vec{z}/\vec{s}]) \equiv \kappa((\varepsilon x.\psi)[\vec{y}/\vec{d}])$;

(3) $[x_i] = [y_i]$ for $i = 1, \ldots, n$,

then $\Sigma \vdash (\varepsilon y.\varphi)[\vec{z}/\vec{s}] = (\varepsilon x.\psi)[\vec{y}/\vec{d}]$ (and similarly for τ). But this is routine using the definition of $[t]$, the (α)–axioms, and the usual syntactic proof that variables bound by quantifiers can be renamed at will.

From here we can complete our definition of a canonical valuation in the same way as above by appealing to Proposition 10.1, and so we can conclude that the following theorem holds.

Theorem 7 (Completeness Theorem). *In each of the following groups of three claims the claims are mutually equivalent.*

(1) $\Sigma \vdash_\varepsilon \varphi$.

(2) $\|\varphi\|^\pi_{\mathcal{M}} = 1$ *in the canonical valuation.*

(3) $\Sigma \models_\varepsilon \varphi$.

(1′) $\Sigma \vdash_\tau \varphi$.

(2′) $\|\varphi\|^\pi_{\mathcal{M}} = 1$ *in the canonical valuation.*

(3′) $\Sigma \models_\tau \varphi$.

2. The Strength of ε and τ

2.1. Introduction.

The effect adding ε to the intuitionistic predicate calculus has on quantifiers is of some philosophical interest, since it was Hilbert's contention that the (ε)-axiom could be used to justify the behaviour of the classical quantifiers in the face of the complaints about classical reasoning raised by the intuitionists. We will see in this section that Hilbert's contention is correct only if we accept principles, like excluded middle (10.4.1),[1] which are intuitionistically unacceptable. This still leaves us with a natural question: which classically valid but intuitionistically invalid principles *does* ε buy for us?

[1] For the remainder of this chapter we will write, for example, (10.4.1) for Proposition 10.4(1).

It is not clear who was the first author to point out that addition of the ε–operator to intuitionistic logic is non-conservative, but this is something that has been remarked by several authors. Most commonly it is pointed out that the formula

$$\exists x.(\exists x.\psi \Rightarrow \psi) \tag{10.5.5}$$

is provable in the first order intuitionistic ε–calculus, while it is not valid in first order intuitionistic logic. This is pointed out, e. g., by [Mints 1974], [Hazen], and [Bell 1993b].

[Bell 1993a] seems to be the first place where it is mentioned that the introduction of ε–terms not only changes the quantifier laws, but also has *propositional* consequences. In that paper he shows that the intuitionistically invalid de Morgan's law

$$\neg(\alpha \wedge \beta) \Rightarrow \neg\alpha \vee \neg\beta \tag{10.4.4}$$

is derivable in the intuitionistic ε–calculus, given that we have constants c and b such that $\forall x.(x = c \vee x \neq c) \wedge b \neq c$ is provable. This perhaps makes clearer why it is an interesting project to investigate the extra deductive strength adding ε–terms confers on intuitionistic logic. For (10.5.5) is merely a rather strange looking quantifier rule, one which, as Hazen remarks, can be a cause of some discomfort to teachers of introductory classes is classical logic. But de Morgan's law is often counted a *logical principle* in the very same introductory courses.[2]

[2] [Bell 1993b] proves that (10.5.5) (and so indirectly that ε) together with the same assumption yields the stronger propositional consequence $(\alpha \Rightarrow \beta) \vee (\beta \Rightarrow \alpha)$ (i. e. (10.4.8)). This strange looking formula has the same unintuitive scent as (10.5.5). However, it is equivalent in intuitionistic propositional logic to $\alpha \Rightarrow (\beta \vee \sigma) \Rightarrow ((\alpha \Rightarrow \beta) \vee (\alpha \Rightarrow \sigma))$ (cf. [Horn p. 398]), which has a more "logical" flavour.

As for the intuitionistic τ–calculus, it seems to have been almost completely uninvestigated. [Bell 1993a] does point out that

$$(10.5.1) \qquad \neg\forall x.\psi \Rightarrow \exists x.\neg\psi$$

can be derived from (τ) in intuitionistic logic, and shows that this formula is independent of the intuitionistic ε–calculus. He concludes from this that "in an existential sense the τ–operator is stronger than the ε–operator" [p. 12]. We will see further evidence for this claim in what follows.

Bell shows in the same paper that while the *law of excluded middle*

$$(10.4.1) \qquad \alpha \vee \neg\alpha$$

is *not* derivable in either the intuitionistic ε–calculus or the intuitionistic τ–calculus, it *is* derivable in the *extensional* intuitionistic ε–calculus, given only that we are working in a theory such that we have two constants c and b such that we can prove $c \neq b$. Bell's independence proofs actually show, as we shall see, that (10.4.1) is not derivable in either the quasi-extensional Hε–calculus or the quasi-extensional Hτ–calculus.

Finally, before proceeding we point out that anything provable in the ε–calculus is provable in the ε^*–, Hε^*–, extensional ε– and H–extensional ε–calculi, and similarly for the τ–case. Almost conversely, a quasi–extensional ε (or τ) structure is also an ε–structure (a τ–structure), and so if we show something independent of the Hε^*–calculus (Hτ^*–calculus) we have also shown it to be independent of the ε– (the τ–) calculus.

2.2. Quantifier Laws.

We will begin by showing that adding ε or τ to intuitionistic logic changes the behaviour of quantifiers.

Theorem 8. *Let $\varphi, \psi \in$ **Wffs** with $x \notin$ **FV**(φ). Then the following formulas are provable in the first order intuitionistic ε–calculus.*

$$(\varphi \Rightarrow \exists x.\psi) \Rightarrow \exists x.(\varphi \Rightarrow \psi) \tag{10.5.4}$$

$$\exists x.(\exists x.\psi \Rightarrow \psi) \tag{10.5.5}$$

Proof. (10.5.5) obviously follows from the (ε)–schema by existential generalization, and (10.5.4) follows by Lemma 10.5. □

Theorem 9. *Let $\varphi, \psi \in$ **Wffs** such that $x \notin$ **FV**(φ). Then we can prove the following formulas in the intuitionistic τ–calculus.*

$$\neg\forall x.\psi \Rightarrow \exists x.\neg\psi \tag{10.5.1}$$

$$(\forall x.\psi \Rightarrow \varphi) \Rightarrow \exists x.(\psi \Rightarrow \varphi) \tag{10.5.2}$$

$$\exists x.(\psi \Rightarrow \forall y.\psi) \tag{10.5.3}$$

$$\forall x.(\varphi \vee \psi) \Rightarrow \varphi \vee \forall x.\psi \tag{10.5.6}$$

Proof. The (τ) axiom gives us $\vdash_\tau \psi[x/\tau x.\psi] \Rightarrow \forall x.\psi$, and so by one of the intuitionistically valid forms of contraposition we have $\vdash_\tau \neg\forall x.\psi \Rightarrow \neg\psi[x/\tau x.\varphi]$, whence

$$\neg\forall x.\psi \vdash_\tau \neg\psi[x/\tau x.\psi]$$

and so by existential generalization

$$\neg\forall x.\psi \vdash_\tau \exists x.\neg\psi.$$

(10.5.1) follows by the deduction theorem.

Next, we know

$$\forall x.(\varphi \vee \psi) \vdash_\tau (\varphi \vee \psi)[x/\tau x.\psi],$$

so if x is not free in φ,

$$\forall x.(\varphi \vee \psi) \vdash_\tau \varphi \vee \psi[x/\tau x.\psi],$$

so by the (τ) axiom

$$\forall x.(\varphi \vee \psi) \vdash_\tau \varphi \vee \forall x.\psi$$

and (10.5.6) follows by the deduction theorem.

Finally, by the τ–axiom $\psi[x/\tau x.\psi] \Rightarrow \forall x.\psi$, and since x is not free in $\forall x.\psi$, mere syntactic congruence tells us $\vdash_\tau (\psi \Rightarrow \forall x.\psi)[x/\tau x.\psi]$. So (10.5.3) follows by existential generalization. So (10.5.2) is also provable, by Lemma 10.5. □

As mentioned, (10.5.5) has often been pointed to as an example that shows that the addition of ε–terms to first order intuitionistic logic is non–conservative. The proof of (10.5.3) is obviously just the corresponding claim for τ. The proof that (10.5.1) is provable in the τ–calculus is from [Bell 1993a]. The other claims, though rather obvious, do not seem to have been remarked upon in the literature.

[Bell 1993a] shows that Markov's principle

$$\forall x.[\neg\neg\beta \Rightarrow \beta] \Rightarrow [\neg\forall x.\beta \Rightarrow \exists x.\neg\beta]$$

is derivable in the intuitionistic ε-calculus. This is an instance of a more general fact, namely that for a *decidable predicate* φ the term $\varepsilon x.\neg\varphi$ functions like $\tau x.\varphi$. This claim has an obvious analogue in the τ-calculus, which turns out to hold too.

Theorem 10. *For* $\varphi \in$ **Wffs**,

(1) $\vdash_\varepsilon \forall x.[\neg\neg\varphi \Rightarrow \varphi] \Rightarrow (\varphi[x/\varepsilon x.\neg\varphi] \Rightarrow \forall x.\varphi)$.

(2) $\vdash_\tau \forall x.[\neg\neg\varphi \Rightarrow \varphi] \Rightarrow (\exists x.\varphi \Rightarrow \varphi[x/\tau x.\neg\varphi])$.

Proof. (1) is proved, unremarked, by Bell as part of his proof of Markov's principle. Since $\neg\alpha \Rightarrow \exists x.\neg\alpha$ is an axiom, the ε-axiom yields

$$\neg\alpha \vdash_\varepsilon \neg\alpha[x/\varepsilon x.\neg\alpha],$$

so

$$\neg\neg\alpha[x/\varepsilon x.\neg\alpha] \vdash_\varepsilon \neg\neg\alpha,$$

and so, since $\alpha \Rightarrow \neg\neg\alpha$ is an intuitionistically valid principle,

$$\forall x.[\neg\neg\alpha \Rightarrow \alpha],\ \alpha[x/\varepsilon x.\neg\alpha] \vdash_\varepsilon \alpha,$$

therefore by universal generalization

$$\forall x.[\neg\neg\alpha \Rightarrow \alpha], \alpha[x/\varepsilon x.\neg\alpha] \vdash_\varepsilon \forall x.\alpha$$

and the result follows by the Deduction Theorem twice.

To prove (2), first note that it is a valid intuitionistic principle that $(\forall x.\neg\varphi \Rightarrow \neg\exists x.\varphi)$. This together with the (τ) axiom gives us

$$\vdash_\tau \neg\varphi[x/\tau x.\neg\varphi] \Rightarrow \neg\exists x.\varphi,$$

so by an intuitionistically valid form of contraposition

$$\neg\neg\exists x.\varphi \Rightarrow \neg\neg\varphi[x/\tau x.\neg\varphi].$$

So, again relying on the fact that $\alpha \Rightarrow \neg\neg\alpha$ is a valid intuitionistic principle, we have

$$\forall x.[\neg\neg\varphi \Rightarrow \varphi], \exists x.\varphi \vdash_\tau \varphi[x/\tau x.\neg\varphi],$$

and the result follows by two applications of the Deduction Theorem. □

Immediately from Theorems 8, 9, and 10 we get

Corollary. *Let* $\varphi, \psi \in$ **Wffs**, *with* $x \notin$ **FV**(φ). *Then*

(1) $\vdash_\varepsilon \forall x.[\neg\neg\psi \Rightarrow \psi] \Rightarrow [\neg\forall x.\varphi \Rightarrow \exists x.\neg\varphi]$.

(2) $\vdash_\varepsilon \forall x.[\neg\neg\psi \Rightarrow \psi] \Rightarrow [\forall x.(\varphi \vee \psi) \Rightarrow (\varphi \vee \forall x.\psi)]$.

(3) $\vdash_\varepsilon \forall x.[\neg\neg\psi \Rightarrow \psi] \Rightarrow \exists x.(\psi \Rightarrow \forall x.\varphi)$.

(4) $\vdash_\varepsilon \forall x.[\neg\neg\psi \Rightarrow \psi] \Rightarrow [(\forall x.\psi \Rightarrow \varphi) \Rightarrow \exists x.(\psi \Rightarrow \varphi)]$.

(5) $\vdash_\tau \forall x.[\neg\neg\psi \Rightarrow \psi] \Rightarrow [(\varphi \Rightarrow \exists x.\psi) \Rightarrow \exists x.(\varphi \Rightarrow \psi)]$.

(6) $\vdash_\tau \forall x.[\neg\neg\psi \Rightarrow \psi] \Rightarrow \exists x.(\exists x.\psi \Rightarrow \psi)$. □

2.3. Propositional Consequences.

As we have already had occasion to mention, one of the most interesting points made in [Bell 1993a] is that the ε–operator changes not only the quantifier rules in intuitionistic logic, but it also (given some modest further assumptions) has propositional consequences. In particular, he shows that with those modest further assumptions any sentence of form (10.4.4) is provable in the intuitionistic ε–calculus. In [1993b] he includes some remarks which show that (10.5.5) implies, given these same modest assumptions, that (10.4.7) is provable, which, given Theorem 8, implies that, on those same assumptions, (10.4.7) is provable in the intuitionistic ε-calculus. Bell also mentioned in conversation that (10.4.7) can be proved from (10.5.3), i. e. the second part of the following theorem holds. For the first part we give a direct proof which is a straightforward adaptation of the proof in [Bell 1993b].

Theorem 11. *If Σ is a set of sentences such that for some $c, b \in$ **Con**, $\Sigma \vdash_\varepsilon \forall x.(x = c \vee x \neq c) \wedge c \neq b$, then for any $\alpha, \gamma \in$ **Wffs**,*

$$\Sigma \vdash_\varepsilon (\beta \Rightarrow \gamma) \vee (\gamma \Rightarrow \beta).$$

If Σ meets these conditions for the τ-calculus,

$$\Sigma \vdash_\tau (\beta \Rightarrow \gamma) \vee (\gamma \Rightarrow \beta).$$

So (10.4.4), (10.4.5) and (10.4.6) are all derivable from Σ in the intuitionistic ε-calculus and in the intuitionistic τ-calculus.

Proof. For both the ε and τ case we begin by choosing a variable x not free in either β or γ and define

$$\alpha \equiv (x = c \wedge \beta) \vee (x \neq c \wedge \gamma).$$

We take the ε-case first. Clearly $\Sigma \vdash_\varepsilon \alpha[x/c] \Leftrightarrow \beta$ and $\Sigma \vdash_\varepsilon \alpha[x/b] \Leftrightarrow \gamma$, so $\Sigma \vdash_\varepsilon \beta \vee \gamma \Leftrightarrow \exists x.\alpha$. Brute syntax tells us that

$$\alpha[x/\varepsilon x.\alpha] \equiv (\varepsilon x.\alpha = c \wedge \beta) \vee (\varepsilon x.\alpha \neq c \wedge \gamma),$$

so $\varepsilon x.\alpha = c, \alpha[x/\varepsilon x.\alpha] \vdash_\varepsilon \beta$ and $\varepsilon x.\alpha \neq c, \alpha[x/\varepsilon x.\alpha] \vdash_\varepsilon \gamma$. Now the (ε) schema gives us $\vdash_\varepsilon \exists x.\alpha \Rightarrow \alpha[x/\varepsilon x.\alpha]$, so

$$\begin{aligned}
\Sigma \vdash_\varepsilon\ & \beta \vee \gamma \Rightarrow \alpha[x/\varepsilon x.\alpha] \\
\vdash_\varepsilon\ & (\beta \vee \gamma) \Rightarrow \alpha[x/\varepsilon x.\alpha] \wedge (\varepsilon x.\alpha = c \vee \varepsilon x.\alpha \neq c) \\
\vdash_\varepsilon\ & [(\beta \vee \gamma \Rightarrow \alpha[x/\varepsilon x.\alpha]) \wedge \varepsilon x.\alpha = c] \vee [(\beta \vee \gamma \Rightarrow \alpha[x/\varepsilon x.\alpha]) \wedge \varepsilon x.\alpha \neq c] \\
\vdash_\varepsilon\ & [\beta \vee \gamma \Rightarrow (\alpha[x/\varepsilon x.\alpha] \wedge \varepsilon x.\alpha = c)] \vee [\beta \vee \gamma \Rightarrow (\alpha[x/\varepsilon x.\alpha] \wedge \varepsilon x.\alpha \neq c)] \\
\vdash_\varepsilon\ & [\beta \vee \gamma \Rightarrow \beta] \vee [\beta \vee \gamma \Rightarrow \gamma] \\
\vdash_\varepsilon\ & (\gamma \Rightarrow \beta) \vee (\beta \Rightarrow \gamma).
\end{aligned}$$

For the τ–case, we can again begin by noting that $\Sigma \vdash_\tau \alpha[x/c] \Leftrightarrow \beta$ and $\Sigma \vdash_\tau \alpha[x/b] \Leftrightarrow \gamma$, but this time we note that it follows that $\Sigma \vdash_\tau \forall x.\alpha \Leftrightarrow \beta \wedge \gamma$. Since

$$\alpha[x/\tau x.\alpha] \equiv (\tau x.\alpha = c \wedge \beta) \vee (\tau x.\alpha \neq c \wedge \gamma),$$

$\tau x.\alpha = c, \alpha[x/\tau x.\alpha] \vdash_\tau \beta$ and $\tau x.\alpha \neq c, \alpha[x/\tau x.\alpha] \vdash_\tau \gamma$. The (τ)–schema gives us $\vdash_\tau \alpha[x/\tau x.\alpha] \Rightarrow \forall x.\alpha$, so

$$\begin{aligned}
\Sigma \vdash_\tau\ & \alpha[x/\tau x.\alpha] \Rightarrow \beta \wedge \gamma \\
\vdash_\tau\ & \alpha[x/\tau x.\alpha] \Rightarrow (\beta \wedge \gamma) \wedge (\tau x.\alpha = c \vee \tau x.\alpha \neq c) \\
\vdash_\tau\ & [(\alpha[x/\tau x.\alpha] \Rightarrow (\beta \wedge \gamma)) \wedge \tau x.\alpha = c] \vee [(\alpha[x/\tau x.\alpha] \Rightarrow (\beta \wedge \gamma)) \wedge \tau x.\alpha \neq c] \\
\vdash_\tau\ & [(\alpha[x/\tau x.\alpha] \wedge \tau x.\alpha = c) \Rightarrow (\beta \wedge \gamma)] \vee [(\alpha[x/\tau x.\alpha] \wedge \tau x.\alpha \neq c) \Rightarrow (\beta \wedge \gamma)] \\
\vdash_\tau\ & (\beta \Rightarrow (\beta \wedge \gamma)) \vee (\gamma \Rightarrow (\beta \wedge \gamma)) \\
\vdash_\tau\ & (\beta \Rightarrow \gamma) \vee (\gamma \Rightarrow \beta).
\end{aligned}$$

The rest of the theorem is immediate from Lemma 10.3. □

Both conjuncts of the assumption above seem to be necessary for us to be able to derive (10.4.7) in each of the ε and τ cases. And both seem to be required to derive (10.4.4) and its equivalents, too, for the ε-case. But in the case of τ we can drop the second conjunct.

Theorem 12. *Suppose for a set of sentences Σ there is some constant c such that $\Sigma \vdash_\tau \forall x.(x = c \vee x \neq c)$. Then for any $\beta, \gamma \in$ **Wffs**,*

$$\Sigma \vdash_\tau \neg(\beta \wedge \gamma) \Rightarrow \neg\beta \vee \neg\gamma.$$

So (10.4.5), and (10.4.6) are also provable for all β and γ.

Proof. Choose a variable y such that $y \notin \mathbf{FV}(\beta)$ and $y \notin \mathbf{FV}(\gamma)$. Define

$$\alpha \equiv (y = c \wedge \beta) \vee (y \neq c \wedge \gamma).$$

Then $\Sigma \vdash_\tau \alpha[y/\tau y.\alpha] \Leftrightarrow \forall y.\alpha \Leftrightarrow \beta \wedge \gamma$. So

$$\Sigma \vdash_\tau \neg\alpha[y/\tau y.\alpha] \Leftrightarrow \neg\forall y.\alpha \Leftrightarrow \neg(\beta \wedge \gamma).$$

So

$$\begin{aligned} \Sigma, \neg(\beta \wedge \gamma) &\vdash_\tau \neg[(\tau y.\alpha = c \wedge \beta) \vee (\tau y.\alpha \neq c \wedge \gamma)] \\ &\vdash_\tau \neg(\tau y.\alpha = c \wedge \beta) \wedge \neg(\tau y.\alpha \neq c \wedge \gamma) \\ &\vdash_\tau (\tau y.\alpha = c \Rightarrow \neg\beta) \wedge (\tau y.\alpha \neq c \Rightarrow \neg\gamma). \end{aligned}$$

But since $\Sigma \vdash_\tau \tau y.\alpha = c \vee \tau y.\alpha \neq c$, we have $\Sigma, \neg(\beta \wedge \gamma) \vdash_\tau \neg\beta \vee \neg\gamma$, and the result follows. The rest of the theorem then follows by Lemma 10.3. □

In the classical ε–case we saw that the ε–Theorems hold whether we are dealing with the extensional or the non-extensional case, so the addition of the extensionality axioms does nothing in classical logic to increase the ε-free consequences of ε–free sets of formulas. The same does not hold true in the intuitionistic case. The following result is one part of showing this.[3]

Theorem 13 (Bell). *Let Σ be a set of sentences such that there are $b, c \in$ **Con** for which $\Sigma \vdash_{Heext} b \neq c$. Then for any $\alpha \in$ **Wffs**,*

$$\Sigma \vdash_{Heext} \alpha \vee \neg\alpha.$$

Proof. Choose a variable $y \notin \mathbf{FV}(\alpha)$ and define

$$\beta \equiv y = b \vee \alpha$$

$$\gamma \equiv y = c \vee \alpha.$$

Obviously $\Sigma \vdash_{Heext} \exists y.\beta$ and $\Sigma \vdash_{Heext} \exists y.\gamma$, so

$$\Sigma \vdash_{Heext} \beta[y/\varepsilon y.\beta] \wedge \gamma[y/\varepsilon y.\gamma],$$

which is to say

$$\Sigma \vdash_{Heext} [\varepsilon y.\beta = b \vee \alpha] \wedge [\varepsilon y.\gamma = c \vee \alpha].$$

So

$$\Sigma \vdash_{Heext} [\varepsilon y.\beta = b \wedge \varepsilon y.\gamma = c] \vee \alpha,$$

[3] Bell reports [Bell 1993a, fn. 11] that this result was discovered by constructing a "stripped down" version of Diaconescu's famous proof that the axiom of choice implies excluded middle in a topos.

and so

$$\Sigma \vdash_{H\varepsilon ext} [\varepsilon y.\beta \neq \varepsilon y.\gamma] \vee \alpha.$$

But $\varepsilon y.\beta \neq \varepsilon y.\gamma \vdash_{H\varepsilon ext} \neg\forall y.(\beta \Leftrightarrow \gamma)$ because of the (Ack) schema, and so $\varepsilon y.\beta \neq \varepsilon y.\gamma \vdash_{H\varepsilon ext} \neg\alpha$. So $\Sigma \vdash_{H\varepsilon ext} \neg\alpha \vee \alpha$. □

2.4. Improving the semantics for τ.

We suggested in the introduction to this chapter that we might be able to improve the semantics above. We will now show one respect in which we can do so, employing some of the results from [Horn]. We begin with some terminology. A *chain* is a linearly ordered Heyting algebra. A *linear algebra* (or *L–algebra*) is a Heyting algebra L such that for all $a, b \in L$, $(a \Rightarrow b) \vee (b \Rightarrow a) = 1$. Clearly every chain is a linear algebra, since in any Heyting algebra $a \Rightarrow b = 1 \iff a \leqslant b$. However, the converse is not true. But

Proposition 1. *If F is a prime filter in an L–algebra Ω, Ω/F is a chain.* □

In a chain $a \Rightarrow b = 1$ if $a \leqslant b$, and $a \Rightarrow b = b$ if $b < a$. So $a^* = a \Rightarrow 0 = 1$ if $a = 0$, and $a^* = 0$ otherwise.

Theorem 11 tells us that the Lindenbaum algebra for any theory in the intuitionistic ε–calculus or τ–calculus which has as a theorem

$$(\dagger) \qquad \forall x.(x = c \vee x \neq c) \wedge c \neq b$$

is an L–algebra. This might suggest that we can modify our semantics above by saying that if we restrict attention to theories of the appropriate sort we can insist that the lattice of truth-values Ω must be an L–algebra. But things are not so simple, for Horn has shown

Proposition 2. *The normal completion of an L–algebra need not be an L–algebra.* □

While we know that the Lindenbaum algebra of a theory in the ε–calculus or the τ–calculus will be an L–algebra (on the assumption that (†) is a theorem), the Lindenbaum algebra need not be complete. So unless we are willing to modify our definition of structure to allow Ω to be non–complete, we cannot get a completeness proof if we restrict attention to L–algebras.[4]

However, since the normal completion of a chain is a chain, no similar problem will result in that case. Let *chain logic* be the system of logic which results if we begin with the first order intuitionistic predicate calculus, then add (10.4.7) to the list of propositional axioms and (10.5.6) to the list of quantifier axioms. Let *chain semantics* be the semantics which results if we insist that Ω must be a chain in the definition of an intuitionistic $\mathcal{L}'$-structure. Horn shows

Proposition 3. *Chain logic is sound and complete for chain semantics.* □

Briefly sketching the method Horn uses in his proof will allow us to see an easy way to put this result to use.[5] Let $\Omega(\Sigma)$ be the Lindenbaum algebra for Σ. A *Q-homomorphism* $h(\Omega(\Sigma)) \to \Omega$ is a Heyting algebra homomorphism such that

$$\bigvee_{t \in \mathbf{Terms}} h(\|\alpha[x/t]\|) = h(\|\exists x.\alpha\|)$$

and

$$\bigwedge_{t \in \mathbf{Terms}} h(\|\alpha[x/t]\|) = h(\|\forall x.\alpha\|).$$

[4] Unless, of course, we can cook up some other method of completing an L–algebra which leaves the result an L–algebra. I can see no way to do this.

[5] This discussion assumes, for simplicity, that $\mathcal{L}$ is countable.

Horn proves

Proposition 4. *If* $1 \neq \|\alpha\| \in \Omega(\Sigma)$, *then there is a prime filter* P *in* $\Omega(\Sigma)$ *such that* $\|\alpha\| \notin P$ *and such that the map* h *defined by* $h(\|\varphi\|) \to \|\alpha\| / P$ *is a* Q*-homomorphism* $h : \Omega(\Sigma) \to \Omega(\Sigma)/P$. □

Horn proves the completeness part of Proposition 3 by arguing as follows. Suppose α is valid in every chain-semantics structure on which all members of Σ are valid. If $\Sigma \nvdash \alpha$, then $\|\alpha\| \neq 1$, so on the homomorphism we get from Proposition 4 $h(\|\alpha\|) \neq 1$ in $\Omega(\Sigma)/P$. But $\Omega(\Sigma)/P$ is a chain by Proposition 1. It is straightforward to generate a chain-valued canonical interpretation on which α is not satisfied but every member of Σ is, thus generating a contradiction.

Now, since $\vdash_\tau \forall x.\varphi \Leftrightarrow \varphi[x/\tau x.\varphi]$ for every φ, the Lindenbaum algebra for Σ in the τ–calculus will be n–meet–attained (with respect to **Wffs**) for all $n \in \omega$, and so if $\Sigma \vdash_\tau \forall x.(x = c \vee x \neq c) \wedge b \neq c$ we can apply Proposition 4 to get a Q-homomorphism onto a chain, and the chain will obviously likewise be n–meet–attained for all $n \in \omega$. So if we rewrite the various definitions of τ–structures above replacing the words 'Heyting algebra' by the word 'chain', we will be able to prove without too much work the completeness for their respective semantics of the various τ–calculi to which we add (†). And since we will still be requiring that the structures have all the necessary meets attained, we will be able to prove soundness by appealing to Horn's result and showing that the (τ)–scheme is valid exactly as we did in §2.

We won't bother to state these results explicitly, since all the redefinitions and restatements would take up a lot of space for little reward.

2.5. Independence Results.

So far we have seen that the addition of either ε or τ to intuitionistic predicate logic is non-conservative in some interesting ways. But there is more to evaluating the extra strength these operators add to our system than just showing that it has certain consequences. We also want to know what cannot be proved in these systems. In particular, we want to know that the addition of these operators does not merely collapse intuitionistic logic into classical logic. This requires independence proofs and so, of course, that we put our semantics to work.

The only semantics in the literature for the intuitionistic ε–calculus is a sound but not complete semantics in [Bell 1993a]. It is not hard to see now why Bell's semantics is sound. His structures are the same as those of Definition 2 except that he does not specify that Ω must be n–join–attained for all n, but he does require that Ω be an inversely well-ordered set with 0. The inverse well-ordering guarantees that each subset of Ω has a maximal element, which guarantees also that *all* joins are attained, so Bell could skip the details of stipulating that a particular subset of all the joins needed to be attained. Since Bell's structures are a subset of our structures, we can borrow his independence results to get this section started. (1) and (4) of the next theorem are from [Bell 1993a].

Theorem 14. *Let $\varphi, \psi \in$ **Wffs** be such that $x \notin \mathbf{FV}(\psi)$. The following schemes are not provable in the intuitionistic ε–calculus.*

(1) $\neg\forall x.\varphi \Rightarrow \exists x.\neg\varphi$.

(2) $\exists x.(\varphi \Rightarrow \forall x.\varphi)$.

(3) $(\forall x.\varphi \Rightarrow \psi) \Rightarrow \exists x.(\varphi \Rightarrow \psi)$.

(4) $\varphi \vee \neg\varphi$.

Proof. Let $\mathcal{L}$ be the language with as its sole non-logical predicate $P \in \mathbf{Pred}_1$. We construct a structure and a valuation such that (1) and (2) are not satisfied. (4) follows, since the law of excluded middle added to intuitionistic logic yields classical logic, and as we have already mentioned, (1) and (2) are classically valid. (3) follows by Lemma 10.4. Let Neg be the set of negative integers and 0, and let $\Omega = \emptyset \cup \{ (\leftarrow, -n] \mid n \in \omega \}$, ordered by inclusion. It is easy to check that this is an inversely well-ordered set with $1 = Neg$ and $0 = \emptyset$. We take as our domain $D = \omega$, and put

$$\mathrm{eq}_{\mathcal{M}}(m,n) = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n. \end{cases}$$

For our choice function we can just let $E(X)$ be the least element of X, and we can interpret P by $\mathcal{R} : \omega \to \Omega$ such that $\mathcal{R}(n) = (\leftarrow, -n]$. $\mathcal{M} = \langle \omega, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, E \rangle$ is easily checked to be a quasi-extensional intuitionistic ε–structure for $\mathcal{L}$. Now let $\varrho : \mathbf{Var} \to \omega$ be defined by $\varrho(v_n) = n$, and let $\varphi \equiv Px$. It is then easy to calculate that if $x \equiv v_m$

$$\|\varphi\|^{\varrho}_{\mathcal{M}} = (\leftarrow, -m]$$

$$(\|\varphi\|^{\varrho}_{\mathcal{M}})^* = (\leftarrow, -m]^* = 0.$$

So

$$\bigwedge_{m \in \omega} \|\varphi\|^{\varrho(x/m)}_{\mathcal{M}} = \bigwedge_{m \in \omega} (\leftarrow, -m] = \emptyset = 0$$

and

$$\bigvee_{m\in\omega}(\|\varphi\|_{\mathcal{M}}^{\varrho(x/m)})^* = \bigvee_{m\in\omega}(\leftarrow,-m]^* = \emptyset = 0,$$

but

$$(\bigwedge_{m\in\omega}\|\varphi\|_{\mathcal{M}}^{\varrho(x/m)})^* = 0^* = 1.$$

Thus

$$(\bigwedge_{m\in\omega}\|\varphi\|_{\mathcal{M}}^{\varrho(x/m)})^* = 1 \not\leq \bigvee_{m\in\omega}(\|\varphi\|_{\mathcal{M}}^{\varrho(x/m)})^*,$$

so $\|\neg\forall x.\varphi \Rightarrow \exists x.\neg\varphi\|_{\mathcal{M}}^{\varrho} \neq 1$.

Similarly, since $\mathcal{R}(n) \neq \emptyset$ for each n, $\|\varphi\|_{\mathcal{M}}^{\varrho(x/n)} \Rightarrow \|\forall x.\varphi\|_{\mathcal{M}}^{\varrho} = 0$ for all n, and so $\bigvee_{n\in\omega}\|\varphi \Rightarrow \forall x.\varphi\|_{\mathcal{M}}^{\varrho(x/n)} = 0 \neq 1$. □

Since it is an easy modification to make $\mathcal{M}$ a model of $\forall x.(x = c \vee x \neq c) \wedge c \neq b$ for $\mathcal{L}$ with the constants c and b added, we have

Corollary. *(Ack) is an indispensable part of the derivation of the law of excluded middle in Theorem 13.* □

Bell comments in [1993a] that his semantics can also be easily modified—by taking well-ordered instead of inversely well-ordered sets as our lattices of truth-values—to get a sound semantics for the intuitionistic τ-calculus. We saw in the preceding section that this is very close to being an adequate semantics for τ.[6] Bell mentions that the independence result (3) of the next theorem can then be proved using that semantics.

[6] Strictly speaking, for the quasi-H-extensional τ-calculus.

Theorem 15. *Let $\varphi, \psi \in$ **Wffs** be such that $x \notin \mathbf{FV}(\psi)$. The following schemes are not derivable in the intuitionistic τ–calculus.*

(1) $\exists x.(\exists x.\varphi \Rightarrow \varphi)$.

(2) $(\varphi \Rightarrow \exists x.\psi) \Rightarrow \exists x.(\varphi \Rightarrow \psi)$.

(3) $\alpha \dot{\vee} \neg\alpha$.

Proof. Again, (1) and (2) are classically valid, and they are equivalent by Lemma 10.5, so it suffices to prove all three if we find a τ–structure $\mathcal{M}$ and a valuation ϱ such that

$$(\dagger) \qquad \|\exists x.(\exists x.\varphi \Rightarrow \varphi)\|^{\varrho}_{\mathcal{M}} \neq 1.$$

We will use the same language as in the proof of Theorem 14, and we will let D, $\mathrm{eq}_{\mathcal{M}}$, E, and ϱ be defined as in that proof as well. But now we will take $\Omega = \omega + 2$, that is Ω is the chain that results if we add a new top element to the chain of all ordinals $\leqslant \omega$. This is obviously a well-ordered set, so the model which results from taking it as our lattice of truth-values will be n–meet–attained for all n. Obviously, 0=0 and $\omega + 1 = 1$.

We interpret Px by the function $\mathcal{R} : \omega \to \omega + 2$ defined by $\mathcal{R}(n) = n$ for all $n \in \omega$. Then

$$\bigvee_{n \in \omega} \|Px\|^{\varrho(x/n)}_{\mathcal{M}} = \omega,$$

and, for each $n \in \omega$, $\omega \Rightarrow n = n$. So

$$\|\exists x.(\exists x.Px \Rightarrow Px)\|^{\varrho}_{\mathcal{M}} = \omega \neq 1. \quad \square$$

In his semantics for the ε–calculus Bell stipulates that that Ω must be an inversely well-ordered set to ensure that each subset of Ω will have a maximal element, thus avoiding the complexity of spelling out that certain subsets have a maximal element. While we saw in §2.4 that there is good reason for saying that the corresponding semantics for τ is very close to adequate, we cannot do so for ε, because we have not proved (10.5.6) in the intuitionistic ε–calculus, and this formula is valid in chains. Indeed, I make the following

Conjecture. *The schema* $\forall x.(\varphi \vee \beta) \Rightarrow (\varphi \vee \forall x.\beta)$, *where we assume that* $x \notin \mathbf{FV}(\varphi)$, *is not valid in the intuitionistic ε–calculus.*

Unfortunately I have so far been unable to prove that any structure in which that schema fails is n–join–attained for all n, so the status of (10.5.6) in the intuitionistic ε–calculus remains an open problem.

2.6. Intuitionistic ε–Theorems.

We mentioned above that there is another focus in the literature on the intuitionistic ε–calculus, one in which the main concern is to investigate what must be done to *eliminate* the non-conservativeness of the ε–operator in intuitionistic logic, thus making the ε–theorems provable. Since this can be regarded as another way of investigating the question of just what the extra strength is that ε adds to intuitionistic logic we will briefly discuss a couple of such results here.

Grigorii Mints has been the most active author working on this subject. His methods are proof-theoretic and so quite different from those used here. But, as in the classical case discussed in Chapter 7, it is sometimes turns out that rather complicated and difficult proof-theoretic results can be given a (to my mind) more perspicuous proof by using semantical methods (so substituting theft for honest toil in the minds of those with other tastes). Mints also works, as is usual for those who take his approach, in a sequent calculus. We will present a couple of his results, translating them into our terms, and will indicate how one of them can be proved using our earlier semantics.[7]

We will also briefly look at another result proved by Helena Rasiowa in [Rasiowa 1956]. This paper presents the first semantical proof of the ε–theorems for classical logic in the literature. Rasiowa's proof is not actually stated for an ε–language. Instead she extends an ordinary first order language $\mathcal{L}_0$ by adding a new n–ary function symbol for each existential formula with n free variables, calling the resulting language $\mathcal{L}_1$, and repeating this procedure ω–many times. But the translation of her proof to an ε–language is straightforward. After presenting her proof of these theorems Rasiowa remarks that "In an analogical way one can prove that the second ε–theorem holds also for some non-classical theories ... in particular for theories based on ... intuitionistic logic." On the face of it this looks rather as though something has gone quite wrong, for it appears to contradict some of the results from earlier in this section which show that ε is non-conservative. So it will be worthwhile to look closely at what Rasiowa has shown to see how these two facts can be reconciled. It turns out that while what she proves in the classical case is recognizably an ε–theorem, in the intuitionistic case it seems at best to be a distant

[7] We present only a couple of Mints' results since most of his work involves extending these results to particular theories such as, e. g., Heyting arithmetic. Mints himself discusses similar results proved by other authors in [Mints 1974, 1990].

cousin.

We will first state the main result of [Mints 1974] in the language of our presentation. Mints proves his result for the *pure intuitionistic predicate calculus without identity*. We then add the ε–terms to this calculus as usual (i. e. as in Chapter 5), but instead of defining a derivation in the ε–calculus as usual, and suffering the non-conservative consequences, we call those sequences of formulas *quasi-derivations*. We designate a special subclass of the set of quasi-derivations the (genuine) *derivations* by insisting that a quasi-derivation is a derivation if and only if any formula in the derivation in which an ε–term $\varepsilon x.\varphi$ occurs must be preceded in the sequence of formulas making up the derivation by the formula $\forall x_1 \ldots \forall x_n \exists x.\varphi$, where $\mathbf{FV}(\exists x.\varphi) = \{x_1, \ldots, x_n\}$.

Let Σ be a set of formulas, φ a formula, $\mathcal{L}'$ an ε-free, identity free language, and $\mathcal{L}$ the corresponding identity free ε–language. We will write '$\Sigma \vdash_M \varphi$' if φ is derivable from Σ in Mints' special sense. We will write '$\Sigma \vdash \varphi$' if φ is derivable from Σ in the identity free intuitionistic predicate calculus.

Theorem 17 (Mints). *If φ is ε-free, $\vdash_M \varphi \Rightarrow \vdash \varphi$.*

Proof sketch. Mints' proof is for the pure intuitionistic ε–calculus, which is *witnessed* by Proposition 10.3. So we can define a *Mints structure* by adding to the definition of an intuitionistic structure clauses which ensure that if $\vdash \exists x.\varphi$, $[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}} \in \{\, [t]^{\varrho}_{\mathcal{M}} \mid \vdash \varphi[x/t] \,\}$, and let $[\varepsilon x.\varphi]^{\varrho}_{\mathcal{M}}$ be an arbitrary member of D otherwise. If we call the Mints calculus the calculus which results if we add to the axioms of intuitionistic predicate calculus only the instances $\exists x.\varphi \Rightarrow \varphi[x/\varepsilon x.\varphi]$ of the (ε)–schema for which $\vdash \exists x.\varphi$, it is clear that we will be able to prove a Soundness Theorem for the Mints calculus and the Mints structures. But since we can construct Mints structures out of any intuitionistic structure and the interpretation of φ is independent of the interpretation of the ε–terms in the Mints structures, the validity of φ in every Mints structure implies the validity of φ in every intuitionistic structure. So $\vdash \varphi$ follows by the completeness of the intuitionistic predicate calculus for these structures. □

Mints makes the restriction to the identity free calculus because, even with his restricted notion of derivation, addition of ε is non-conservative in the version of the intuitionistic predicate calculus with identity he has in mind. This is because Mints seems to be considering a calculus in which the Substitutivity of Identicals is taken to be an axiom scheme, and so (A6) is valid. And in this case we have the following example, which Mints credits to H. Osswald.

Proposition 5. *If we extend the Mints calculu by adding the symbol and the axioms for identity, including (A6), we can prove the following formula, which is not provable in the intuitionistic predicate calculus with identity.*

$$[\forall x.\exists y.((x = 0 \land y = 1) \lor (x = 1 \land y = 1) \lor (x = c \land y = c)) \land 0 \neq 1]$$
$$\Rightarrow (c = 0 \lor c \neq 0). \quad \square$$

Mints' proof of this makes very clear how an extensionality assumption such as (A6) can result in non-conservativeness. Mints' proof involves putting $[(x = 0 \land y = 1) \lor (x = 1 \land y = 1) \lor (x = c \land y = c)] \equiv Pxy$, then running through the cases for

x in $\varepsilon y.Pxy$. In all the cases except where $\varepsilon y.P0y = 1$ and $\varepsilon y.Pcy = c$, without appealing to (A6), we can either prove $c = 0$ or on the assumption that $c = 0$ we get a contradiction with $0 \neq 1$. In the last case, assume that $c = 0$. By (A6), $\varepsilon y.P0y = \varepsilon y.Pcy$, so $0 = 1$, and we have a contradiction with $0 \neq 1$ and we can conclude that $c \neq 0$.

Since we have not included the substitutivity of identicals as an axiom scheme, (or to say the same thing another way, since we have distinguished between the ε–calculus and the Hilbertian ε–calculus) this proof does not go through in the version of the intuitionistic ε–calculus we have been working with and so we could have stated Proposition 4 for the calculus with identity.

Finally, the non-conservativeness due to (A6) also depends crucially on the fact that we are allowing $=$ to be interpreted by any map which ensures that the identity axioms are satisfied. This is made clear by one more result from Mints, the proof of which can be found in [Mints 1990].

Proposition 6. *Write* $\vdash_{M=}$ *for derivability in the Mints calculus with identity to which we add the additional identity axiom* $\forall x.\forall y.(x = y \vee x \neq y)$ *(i. e.* $M_=$ *is the Mints calculus with decidable identity). Then if* φ *is an* ε*–free formula,*

$\vdash_{M=} \varphi \implies \vdash \varphi$. $\square$

In effect, Mints solves the "problem" of the non-conservativeness of ε in intuitionistic logic by allowing us to use only ε–terms which we can know in advance won't get us anything new. So while we have officially let all the ε–terms into our language $\mathcal{L}$, we might as well not have since most of them can play no role in the real business of the language—that is, they cannot appear in any derivation. As he points out in [Mints 1990], this amounts in effect to regarding ε as only partially defined.

Rasiowa's version of the ε-theorem turns out to be rather similar in spirit. We shall not prove her result here.[8] We will content ourselves with trying to state it clearly and with indicating why it doesn't contradict anything proved above.

First, Rasiowa gives a proof which was intended in the first instance to apply to classical logic. Let $\mathcal{L}'$ be an ε-free language. Rasiowa's first step is to translate the (proper) axioms Σ of any theory in $\mathcal{L}'$ into equivalent *prenex* formulas (see Definition 7.8). This is always possible in the classical case, but in the intuitionistic case it is not, so Rasiowa merely states her result for the case of theories with all prenex axioms, which constitutes a considerable restriction on the generality of the result. The next step is to follow the procedure described after Definition 7.8 above for constructing the *Skolem resolutions* α^* of each formula $\alpha \in \Sigma$. The language $\mathcal{L}^*$ is the language that results if we add the Skolem functions and Skolem constants that result to the language $\mathcal{L}'$. Write 'Σ^*' to denote $\{\, \alpha^* \mid \alpha \in \Sigma \,\}$. This is another restriction that is made in the intuitionistic case—we no longer repeat the procedure ω many times. What Rasiowa shows is

Proposition 7. *If φ is a quantifier free formula of $\mathcal{L}'$,*

$\Sigma^* \vdash \varphi \iff \Sigma \vdash \varphi$. $\square$

The final restriction, which is added in the *Errata* to [Rasiowa 1956] is the

[8] A clearer version of the proof of [Rasiowa 1956] can be found in [Rasiowa and Sikorski, pp.453-59].

requirement that φ must be quantifier free.

The reason I say this result is similar to Mints' results is that in both cases we are, in some sense, not in fact adding the full complement of ε-terms to the calculus, and this is what allows us to prevent it from strengthening the intuitionistic predicate calculus. Mints officially works with the full ε–language, but in his calculus most ε-terms lead a decidedly second-class existence. Rather than restricting the notion of derivation, Rasiowa restricts the number of ε–terms that get added to $\mathcal{L}^*$—we only add an ε–term for each existential quantifier that shows up in the set of (prenex) axioms Σ. The restrictions on this result should not be underestimated, and the analogy between it and the ε–theorems we can prove for the classical case should not be taken without a grain of salt.

For instance, we could not attempt to add axioms of a form similar to (Ack) to Σ. Instances of (Ack) are not prenex and, clearly, Theorem 10.6.(2) tells us that it is one of the forms of non-prenex formulas that we are not going to be able to transform in general into an equivalent prenex form. Secondly, the conservativeness result is only for *quantifier free* theorems, and so the result does not apply to, for instance, general schemes like de Morgan's law or the law of excluded middle. Finally, the restriction to $\mathcal{L}^*$ is significant because it is, in a sense, doubly restricted compared to ε. First, of course, just as Mints does, Rasiowa only adds a term $\varepsilon x.\varphi$ in cases where $\exists x.\varphi$ is provable, since all of them turn up in prenex *axioms* (and, indeed, for Rasiowa we may only have ε–terms for a small subset of these cases.) The second restriction is that whereas adding an extra clause to the recursive definition of well-formed expression has the effect of carrying out the process of adding new ε-terms infinitely many times. Rasiowa's proof only involves adding those terms generated by a single pass over the axioms. So if, for example, $\exists x.Px$ and $\exists y.Rzy$ are two axioms, Rasiowa's procedure will introduce terms we can refer to as $\varepsilon x.Px$ and $\varepsilon y.Rzy$, but we will not get a term $\varepsilon z.(\varepsilon x.Px = \varepsilon y.Rzy)$. Examination of, for instance, the proof of Theorem 13, shows how important it is to have ε-terms generally available if we hope to prove general results, and so this too is a significant restriction.

CHAPTER XII

TOWARDS A GENERAL THEORY OF TERM FORMING OPERATORS

It was mentioned in Chapter 1 that investigations of the general theory of term forming operators have so far tended to concentrate on extensional term forming operators. While this is harmless enough in the classical case, since the extensionality assumption does not seem to carry with it any particular deductive strength in classical logic, we saw in the last chapter that the assumption is not one that we can make in first order intuitionistic logic with the same equanimity. For the effect of that assumption in the ε-calculus, as we saw, was (essentially) to reduce the intuitionistic case to the classical case. Since, as we also saw, there are perfectly satisfactory intuitionistic models of the ε-calculus, the assumption is hardly a harmless one.

We also saw in Chapter 1 that the reason people have concentrated on the extensional case is that [Corcoran and Herring] raise some objections to the treatment of term forming operators in [Hatcher 1968]. In Hatcher's original semantics the rules of Universal Instantiation, Substitutivity of Identicals, and Existential Generalization do not hold in full generality, and the interpretation of terms depends on the order in which variables occur in a formula in a way which, so Corcoran and Herring claim, is unnatural if the formula is, for example, an equation. Now that we have seen good reason to not concentrate on the extensional case, we will have to regard these not as objections to, but as consequences of the move to a semantics for non-extensional operators.

The discussion in this chapter will be both less formal and more programmatic than that in earlier chapters. For much of it we will be able to proceed by referring to earlier chapters for details. §1 will begin with a very brief sketch what is called the "standard semantics" in [da Costa and Mortensen], a review of the literature on the general theory of term forming operators. I will then briefly describe the lessons I think the earlier chapters of this thesis teach us about this account and the changes that ought to be made to it. §2 will be a brief sketch of a general theory which implements these changes in a much more general theory. Finally, in §3 which, at long last, is the final section of the thesis, I will briefly indicate some of the directions in which the research described in this thesis needs to be further extended.

1. The "Standard Semantics"

In Chapter 5 we described the syntax of $\mathcal{L}$, a language for an arbitrary term forming operator σ, and we defined the extensional σ-calculus to be the calculus which results if we take (σext) and ($\sigma\alpha$) as additional axiom schemes in the first order predicate calculus. The investigations of what has so far been taken to be the general theory of term forming operators in first order logic has, since [Corcoran and Herring], been primarily concerned with investigating the semantics described in [Corcoran, Hatcher, and Herring] (see, e. g. [Hatcher, 1982], [da Costa, 1980]).

This semantics begins with the idea of interpreting σ by an arbitrary function $\mathcal{P}(D) \to D$ in much the manner in which we used a choice function on $\mathcal{P}(D)$ to interpret ϵ in Chapter 7.

Translating into our notation, what Corcoran, Hatcher and Herring do is first of all to define, for any structure $\mathcal{M}$ for $\mathcal{L}'$, the σ-free sublanguage of $\mathcal{L}$, for any ϱ in $\mathcal{M}$, and for $\varphi \in$ **Wffs** and $x \in$ **Var** the *truth-set for φ and x under ϱ*, viz.,

$$M = \{ d \in D \mid \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \}.$$

They next make the following definition.

Definition 1. Let $f : \mathcal{P}(D) \to D$. If $\mathcal{M} = \langle D, \mathcal{R}, \mathbf{c} \rangle$ is a classical **2**-valued structure for $\mathcal{L}'$, then $\mathcal{M} = \langle D, \mathcal{R}, \mathbf{c}, f \rangle$ is a σ-structure for $\mathcal{L}$. We extend the definition of $\|\varphi\|_{\mathcal{M}}^{\varrho} \in \mathbf{2}$ and $[t]_{\mathcal{M}}^{\varrho} \in D$ from the formulas and terms of $\mathcal{L}'$ to all of **Wffs** and **Terms** by rewriting Definition 6.1, except that we write $\mathcal{L}$ in place of $\mathcal{L}'$, and we add the clause

$$[\sigma x.\varphi]_{\mathcal{M}}^{\varrho} = f(M)$$

where M is the truth-set for φ and x under ϱ.

It should come as no surprise to readers of Chapter 7 that Corcoran et al. are then able to prove that this semantics is complete for the extensional σ-calculus.[1] We will write '$\Sigma \vdash_{\sigma ext} \varphi$' for the claim that φ is derivable in the extensional σ-

[1] In [Corcoran, Hatcher, and Herring] they actually use a single more complicated schema to handle both the functions of making the interpretation of the σ-terms extensional and making sure we can rename bound variables in the σ-terms which they call the *truth-set principle*. It was immediately pointed out by da Costa in his review ([da Costa, 1973]) that adding the truth-set principle as an axiom schema is equivalent to adding both the ($\sigma\alpha$) and (σext) schemas. By the time of [Hatcher 1982], Hatcher adopts a strategy more like da Costa's.

calculus, and will temporarily will use '$\models$' to describe the usual semantic relations in the structures just defined.

Theorem 1. *Let* $\varphi \in$ **Wffs** *and* $\Sigma \subseteq$ **Wffs**. *Then* $\Sigma \vdash_{\sigma ext} \varphi \iff \Sigma \models \varphi$.

Proof sketch. We will satisfy ourselves here with merely sketching a proof. The strategy involves two steps. First we consider the Hilbertian σ-calculus.[2] We reduce the soundness and completeness of a semantics for this calculus to the well-known soundness and completeness of the first order predicate calculus with function symbols.

We carry this out as follows. By making the obvious modifications to the definition of *ground terms*, Definition 8.1, (i. e. putting σ for ε) we get a set of terms such that each σ-term of $\mathcal{L}$ is the result of substituting a unique sequence of terms for the free variables of a uniquely determined ground term. We can then construct another language $\mathcal{L}''$ in which we replace each ground term $\gamma(t)$ of $\mathcal{L}$ with n free variables by a new n-place function symbol $f_{\gamma(t)}$. It is then not hard to define a map $r : \mathcal{L} \to \mathcal{L}''$ which is a bijection of terms to terms and formulas to formulas which allows us to prove that:

(1) There is an exact correspondence between proofs in $\mathcal{L}$ and proofs in $\mathcal{L}''$.

We then define a semantics for σ by assigning a function $f_{\gamma(t)} : D^n \to D$ to each ground term $\gamma(t)$ with n free variables, and adding to our definition of interpretation

[2] Corcoran et al. consider the calculus which results if we simply drop the truth-set principle (cf. note 1) from the extensional calculus, but dropping ($\sigma\alpha$) does nothing but add unnecessary complications, and since they adopt the general principle of Substitutivity of Identicals as an axiom scheme and, as was shown in [Asser] for the ϵ-case (cf. the discussion in Chapter 8§1), this is equivalent to starting with our identity axioms and adding the axiom scheme (σA6), they are essentially working with the Hilbertian σ-calculus.

a clause which states that, if $\vec{x} = \langle x_1, \ldots, x_n \rangle$ is the sequence of free variables in $\gamma(t)$ and $t \equiv t'$, where t' is an alphabetic variant of $\gamma(t)[\vec{x}/\vec{t}]$ for a sequence of terms $\vec{t} = \langle t_1, \ldots, t_n \rangle$, then

$$[t]^{\varrho}_{\mathcal{M}} = f_{\gamma(t)}([t_1]^{\varrho}_{\mathcal{M}}, \ldots, [t_n]^{\varrho}_{\mathcal{M}}).$$

Call such a structure a *pseudo-structure*. The same map r now can be used to const. ıct a map of $\mathcal{M}$, a pseudo-structure of this sort for $\mathcal{L}$, into a structure $r(\mathcal{M})$ for $\mathcal{L}''$ such that:

(2) $[t]^{\varrho}_{\mathcal{M}} = [t]^{\varrho}_{r(\mathcal{M})}$ and $\|\varphi\|^{\varrho}_{\mathcal{M}} = \|\varphi\|^{\varrho}_{r(\mathcal{M})}$.

We thus can prove, using (1) and (2), the soundness and completeness of the classical Hilbertian σ-calculus for the quasi-structures.

The second step is to show that the pseudo-structures on which every instance of (σext) is satisfied are precisely the structures of Definition 1. The result then obviously follows from the soundness and completeness theorems for the pseudo-structures. But it is clear that we can define, for any structure, a pseudo-structure which gives precisely the same interpretations to terms and formulas. To go in the other direction, we can proceed as we did in defining our choice function E in Chapter 7, making the obvious modifications. Once again the extensionality principle guarantees that f will be a function of the appropriate sort. □

2-valued extensional semantics of this sort are sometimes called *standard semantics*. Such semantics have been investigated in some detail by da Costa and his collaborators (cf. [da Costa 1980] and [da Costa and Druck], where it is shown that many of the results of standard first order model theory can be proved for calculi with extensional term forming operators.) The *generality* of this analysis consists in the fact that we can make sure that the axioms which characterize a particular operator are satisfied by placing a restriction on the kinds of function $f : \mathcal{P}(D) \rightarrow D$ we allow to interpret σ. So if we insist that f be a choice function which assigns an arbitrary element of D to $\emptyset$, we get the semantics for ε described in Chapter 7. Da Costa uses this fact to generate the model theoretic results for ε mentioned at the end of Chapter 7. If we want to have σ function as a description operator, we can insist that f act like a choice function on singletons, but that it assign a fixed but arbitrary $d \in D$ to non-singletons (if we add the usual additional axiom to make this last clause appropriate).

However, while this semantics is general in the sense that it allows us to define different extensional term forming operators at our pleasure by restricting the function f as we see fit, limited only by our imagination and the hope that the result will be somehow useful, it is in a variety of ways not general enough. Most of the respects in which it is not appropriate have already been discussed in some guise earlier in this thesis, so we will be brief here, and will allow our discussion to be a trifle imprecise to avoid getting bogged down in details.

First of all, it is not appropriate to simply insist that the σ-terms should be interpreted strictly by application of f to the truth-set for a formula and a variable (in a structure under a valuation). In [Corcoran and Herring], the paper where they originally argue for this approach, it is suggested that this is appropriate because "common sense seems to dictate that speculation concerning a general semantic analysis of [term forming operators] should be preceded by an examination of 'standard' [operators]," and that "the few [operators] in common usage" share that feature [pp. 648–52]. Now, Hilbert's τ operator does not show up on their list of operators in common usage, and perhaps it is not in common usage, but it is one of the few operators which is at all well known whose use is not confined to any "special science" (e. g. set theory or number theory), so a general theory should account for it. In the 2-valued case we can interpret τ by using what we might call an *anti-choice function*, i. e. a function such that

$$f(X) \in D - X$$

for $X \subseteq D$ and such that, for an arbitrary $d \in D$, $f(D) = d$. But it is rather strange to think of an anti-choice function as really being a map *from the truth-set*. And, more importantly, this is obviously going to lead us into troubles in any non-2-valued case.

This brings us to the second problem. The straightforward formulation in terms of truth-sets given by Corcoran, Hatcher and Herring is only a plausible analysis if we restrict attention to 2-valued semantics. This point is discussed in Chapter 9 for the case of ε. For if we want to satisfy the ε axiom, we will need to have the interpretation depend on something other than the truth-set, for the truth-set is too often empty. And now our earlier worries about τ become acute, for we want to have the interpretation of $\tau x.\varphi$ be a member of our domain which satisfies φ *to the least possible extent*, which will no longer be something that holds for arbitrary members of the complement of the truth-set for φ. And, of course, in a non-2-valued semantics we actually need a stronger extensionality principle than (σext) if we are going to get a completeness proof, as was demonstrated in the cases of ε and τ in Chapters 9 and 11.

Thirdly, as was briefly discussed in Chapter 1, and as was demonstrated by the results of Chapter 11, the insistence on extensionality is not a natural one for operators like ε. It is also certainly not a harmless one because, for instance, it (roughly speaking) collapses the intuitionistic ε-calculus into the classical ε-calculus. So any general treatment of term forming operators ought to be able to handle non-extensional term forming operators, since the non-extensional intuitionistic ε-calculus does not imply excluded middle, and it has perfectly good models.

Finally, we note that we will see presently that one of the key complaints raised by Corcoran and Herring in their attack on Hatcher, namely that his semantics leaves open the possibility that the order in which things occur in a term can be relevant to its interpretation when we would normally think it should be irrelevant, while it also holds true for our account, is actually a virtue because it allows us to define certain operators which *rely* on this fact for their peculiar usefulness. We will also have something to say below about how we can ameliorate these problems in cases where they are genuinely counter-intuitive.

2. A Sketch of a New Framework

The preceding discussion suggests that we should aim at a general semantics which satisfies the following two desiderata. Simply by virtue of the nature of the formation rules for our σ-terms, it seems sensible to insist that the interpretation of

$\sigma x.\varphi$ (for ϱ in a particular $\mathcal{M}$) should somehow depend on the interpretation of φ with respect to the valuations $\varrho(x/d)$. But this can't be simply a dependence on the truth-set, since we want to allow the *characteristic axioms* which are the primary vehicles for giving these operators separate identities to take whatever form we can cook up that might be useful. So we will want our general semantics to make the interpretation of $\sigma x.\varphi$ depend on the interpretation of φ in whatever way is required by the characteristic axioms for σ. The second desideratum is that our new framework should be flexible enough to allow us to handle both extensional and non-extensional operators. Indeed, it would be best if it could handle a wide variety of different *degrees of extensionality*.

We will begin with the second problem. We have so far dealt with a few different degrees of extensionality. In Chapters 7 and 8, in particular, we dealt with the classical-extensional, classical-Hilbertian and classical ε-calculi, for the 2-valued case. And the way we dealt with giving the varying degrees of extensionality in the two less extensional cases was to have the interpretation of $\varepsilon x.\varphi$ depend not only on the truth-set for φ and x, but also on the syntactical form of $\varepsilon x.\varphi$. Specifically, we made it depend on the ground term or skeleton of $\varepsilon x.\varphi$, and on the interpretations of a particular set of terms. This was possible because we had been careful to ensure that the definitions of *ground terms* and *skeleton terms* were such that each ε-term could be obtained from one and only one ground term (skeleton term) by simultaneous substitution of a uniquely determined sequence of terms[3] for the free variables of the ground term (skeleton term) and renaming of bound variables.

[3] Up to renaming of bound variables, that is.

In effect, what we did was to divide the set of ε terms into what we might call *syntactical congruence classes*[4] and to make the interpretation of each ε term $\varepsilon x.\varphi$ depend not only on the interpretation of φ, but also on the syntactical congruence class to which $\varepsilon x.\varphi$ belongs.

There is no reason at all to suppose that this approach will only work for the few simple cases we have worked with so far. There are many ways we might characterize the possible divisions of the σ-terms into appropriate congruence classes. The one given by the next proposition will be useful. We want the division to be into classes which are mutually exclusive and jointly exhaustive. If R is a relation, we will say (as we did in Chapter 3) that a term t is a *normal form* with respect to R if, for each term t', Rtt' implies $t \equiv t'$. We say that t' *has a normal form* if there is a normal form t such that $Rt't$.

Proposition 1. *The set of σ-terms is divided into a set of mutually exclusive and jointly exhaustive subsets by a relation $\approx$ if and only if there is some relation R such that:* (1) *R has the Church-Rosser property;* (2) *$\approx$ is the symmetric, transitive closure of R; and* (3) *every σ-term has a normal form with respect to R.*

Proof. Suppose t_1, t_2, t_3 and t_4 are all σ-terms, and suppose that we have defined a relation R, which we might call a *reduction relation*, which has the Church-Rosser property, i. e. if Rt_1t_2 and Rt_1t_3, then there is a t_4 such that Rt_2t_4 and Rt_3t_4. We can then take $\approx$ to be the symmetric, transitive closure of R, and we will be able to prove a Church-Rosser theorem for $\approx$, using exactly the same method as was used

[4] These are, properly speaking, just a sort of equivalence classes. However, we will use this name to avoid suggesting that the relation of logical equivalence needs to be involved.

to prove Theorem 3.1 using Lemma 3.1. A term t' has a normal form if there is a normal form t such that $t \approx t'$, and the corollaries (1) and (2) to Theorem 3.1 holds. That is, each t has at most one normal form, and if t and t' are distinct normal forms, then $t \not\approx t'$. It is not hard to see that if it is also the case that every term has a normal form, then $\approx$ divides the set of terms into a set of mutually exclusive and jointly exhaustive sets of terms.

Conversely, if we have a set of mutually exclusive, jointly exhaustive subsets of the set of σ-terms, $\{ X_i : i \in I \}$, we first choose an arbitrary element t_i^* from each X_i. Define $\approx$ to be the relation such that $t \approx t'$ if and only if t and t' are in the same X_i, and put for R the relation such that for each t in X_i, Rtt' if and only if $t' \equiv t_i^*$. The result is then obvious. □

We are not stuck with such straightforward categorizations of the σ-terms as the skeleton terms and the ground terms. For example, we can use the above proposition to dissolve some of the "problems" with Hatcher's original semantics pointed to by Corcoran and Herring. In particular, we need not adopt the full strength of (σext) to eliminate the possibility that $\sigma x.x = y$ will get a different interpretation from $\sigma x.y = x$, that $\sigma x.(\alpha \wedge \beta)$ will get a different interpretation from $\sigma x.(\beta \wedge \alpha)$, and so on, because we can define a relation with the Church-Rosser property which will put these terms into the same syntactical congruence classes.

Example 1. We might well find it objectionable that our semantics for the classical ε-calculus in Chapter 8 is such that it is possible that $\varepsilon x.(\alpha \wedge \beta)$ gets a different interpretation from $\varepsilon x.(\beta \wedge \alpha)$. Returning to our heuristic device of thinking of $\varepsilon x.\varphi$ as the ideal object with the property φ, we can see some merit in Corcoran and Herring's complaint. What reason could there be for thinking that the ideal fat philosopher, for instance, might be different from the ideal philosopher who is fat? We can try to fix this problem as follows.

For now, let t, t', etc. refer to arbitrary σ-terms of a fixed language $\mathcal{L}$. We can obviously order the well-formed expressions of $\mathcal{L}$ in such a way that less complex expressions come before more complex expressions. We can get the relation we want by aping the definitions in the proof of Lemma 3.1. We can define an appropriate reduction relation on the σ-terms by taking as *redexes* expressions of the following sorts, where $\alpha < \beta$ and $t < t'$ according to the just supposed ordering:

(1) $\sigma x.(t' = t)$

(2) $\sigma x.(\beta \wedge \alpha)$

(3) $\sigma x.(\beta \vee \alpha)$

(4) $\sigma x.\beta$ where this is a free occurrence of a subterm of some σ-term t, t is not a skeleton term, $\sigma x.\beta$ is itself not a skeleton term, and there is no subterm of $\sigma x.\beta$ which is a redex of type (1), (2), or (3).

For redexes of each of these forms the *contractums* are, respectively,

(1) $\sigma x.t = t'$

(2) $\sigma x.(\alpha \wedge \beta)$

(3) $\sigma x.(\alpha \vee \beta)$

(4) $\kappa(\sigma x.\beta)$.

It is straightforward to adapt our earlier definitions to get new definitions of *con-*

traction, *residual*, *minimal redex*, and *minimally complete derivation*. By using considerable patience and very modest ingenuity, we can use the proofs of Chapter 3, especially the proof of Sublemma 3.2, to construct a proof that the relation of *being minimally completely derivable from* has the Church-Rosser property. The relation of reduction that we are after is the transitive closure of being minimally completely derivable from, so by Sublemma 3.3 this relation also is Church-Rosser. And so we can prove a Church-Rosser Theorem for the symmetric, transitive closure of the reduction relation.

The normal forms for this relation are those skeleton terms which are appropriately ordered in the sense that for any conjunction, disjunction or equation which is β in a subterm of form $\sigma x.\beta$, the first conjunct, disjunct or the left side of the equation is less than the second conjunct, disjunct, or the right side of the equation according to our ordering of the well-formed expressions. It is not hard to see that every σ-term has a normal form.

This way of dividing up the σ-terms would obviously be suitable to a σ-calculus to which we add the extensionality axiom schemes:

$$\sigma x.(t = t') = \sigma x.(t' = t)$$

$$\sigma x.(\alpha \wedge \beta) = \sigma x.(\beta \wedge \alpha)$$

$$\sigma x.(\alpha \vee \beta) = \sigma x.(\beta \vee \alpha).$$

All of these are, of course, consequences of (σext), but the second and third are not consequences of (σA6).[5]

[5] Making our interpretations of σ-terms depend on which normal form they have according

Now, examination of the non-extensional semantics given in Chapter 8 will show that the interpretations of the ε-terms did not depend only on the truth-set for the term and the normal form of the term. In these cases we also had our choice function E depend on a sequence of elements of the domain of our interpretation—a sequence with the same length as the number of free variables in the normal form. Which such sequence was relevant to the interpretation of $\varepsilon x.\varphi$ was determined by the interpretation of the terms which must be substituted in the ground (skeleton) term to get $\varepsilon x.\varphi$ back again. If we make our interpretation depend on the normal form as defined in the above example, we will again want to make it also depend on the interpretation of the terms which must be substituted into the normal form to retrieve the original term. In the example above, however, it is not a straightforward matter of substituting terms into the normal form to get the original term back again. We want $\sigma x.(\beta \wedge \alpha)$ to give us the same sequence as $\sigma x.(\alpha \wedge \beta)$, not just the same normal form. So we must take the relevant sequence for $\sigma x.\beta$ to be the sequence of interpretations of the terms which, when substituted simultaneously for the free variables of the normal form, give a term which can be transformed by a series of reversals of the steps (1)—(3) above, and renaming of bound variables, gives us the original term back again.

What this tells us is that if we are going to allow any categorization of the σ-terms meeting the conditions of Proposition 1, then we are going to need a more

to this example would still leave us with interpretations which are rather more intensional that Corcoran and Herring would like, since $\sigma x.c = x$ might still get a different interpretation from $\sigma x.y = x$ even though c and y get the same interpretation. We would need to make adaptations of the sort suggested in the example and work with ground terms instead of skeleton terms to eliminate that "problem".

general formulation of what else the interpretation of the particular σ-terms will depend on than might be suggested by our earlier examples. For what sequence we will want to consider is clearly going to depend crucially on the nature of the relationship between the σ-terms and their normal forms. My suggestion is that we use the following, which is clearly a condition satisfied by the examples we have considered so far.

Stipulation 1. Note that Proposition 1 tells us that we have an onto function $f : S \to N$, where S is the set of σ-terms and N is the set of normal forms. We will now require that f must meet the following ***reversibility condition***.

For any $n \in N$, if n has $\vec{x} = \langle x_1, \ldots, x_m \rangle$ as its sequence of occurrences of free variables, there must be a set of sequences of terms $\{ \langle t_1, \ldots, t_m \rangle_i : i \in I \}$ such that each such sequence is associated with a set

$$(n, \langle t_1, \ldots, t_m \rangle_i) \subseteq f^{-1}[n]$$

in such a way that

$$\bigcup_{i \in I} (n, \langle t_1, \ldots, t_m \rangle_i) = f^{-1}[n]$$

and for $i, j \in I$ with $i \neq j$,

$$(n, \langle t_1, \ldots, t_m \rangle_i) \cap (n, \langle t_1, \ldots, t_m \rangle_j) = \emptyset.$$

We can now begin to see the shape that our general semantic account is going to have to take. If we continue to write $f : S \to N$ for the map taking σ-terms to normal forms, and if we write $(\langle n, \langle t_1, \ldots, t_n \rangle_t \rangle)$ for the subset of $f^{-1}(f(t))$ to which t belongs, then we will want to interpret our σ-terms by some function which, writing S for the set of finite sequences of D,

$$g : \mathcal{P}(D) \times \mathcal{N} \times S \to D.$$

As usual, we will only be concerned with the value of this function in cases where the length of $s \in S$ is the same as the number of occurrences of free variables in $n \in \mathcal{N}$.

Of course, our discussion of the first component of this function is still to come. Before getting on with it, though, we will note a few points that perhaps deserve mention. For those cases in which we do want to treat our operator extensionally, we essentially will want our function g to ignore the last two components. This is easily achieved by taking all the σ-terms to belong to the same syntactical congruence class. As is implicit in what we said above in proving Proposition 1, we can then take any term we choose as our sole normal form for this congruence relation. So if we take $\sigma x.x = x$, then there is only one sequence which has the appropriate number of terms in it, namely the empty sequence, and so we only care about the values of our function g for these values of n and s. And, obviously, there is an isomorphism between the set of functions $g : \mathcal{P}(D) \times \sigma x.x = x \times \{\emptyset\} \to D$ and the set of functions $\mathcal{P}(D) \to D$.

The second point to note about this is that this will give us only a semantics for the calculus with the *rule form* of the extensionality condition, which is to say it will be a semantics for the *quasi-extensional* σ-calculus. We can get a semantics for the calculus with the relevant extensionality *axiom* just as we did in Chapters

9 and 11, namely by taking the quasi-extensional structures to be quasi-structures and defining an actual structure to be a punctilious quasi-structure.

Finally, we note that the adaptations we needed to make to (σext) and (σ^*) in order to handle the non-**2**-valued cases of ε and τ are clearly going to need to be adopted more generally. More specifically, we will want to adopt as our extensionality condition not (σext) but, where for $\sigma x.\beta$ we write '$Ax(\sigma x.\beta)$' for the instance of the characteristic axiom scheme in which $\sigma x.\beta$ occurs except that $\sigma x.\beta$ has been replaced by x,

$$\forall x.(Ax(\sigma x.\varphi) \Leftrightarrow Ax(\sigma x.\psi)) \Rightarrow \sigma x.\varphi = \sigma x.\psi.$$

The modification needed to get the rule version of this condition is obvious.

The final step we must take is that we must clarify the nature of the dependence of the function g on the first component. The proposal I have to make for this is an obvious modification of the "truth-set" proposal of Corcoran and Herring. For since we are here considering operators which can be defined by axioms, rather than using the truth-set for the specific formula β when interpreting $\sigma x.\beta$, we must take instead the truth-set for the *characteristic axiom* for the operator in question. This is, in effect, what we have done in the cases of ε and τ, since

$$\{\, d \in D \mid \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} = \bigvee_{d \in D} \|\varphi\|_{\mathcal{M}}^{\varrho(x/d)} \,\} = \{\, d \in D \mid \|\exists x.\varphi \Rightarrow \varphi\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \,\},$$

and similarly, *mutatis mutandis*, for τ. In other words, we will take the *satisfaction set* for $\sigma x.\varphi$ in $\mathcal{M}$ under ϱ to be the set of $d \in D$ which make $Ax(\sigma x.\varphi)$ true under $\varrho(x/d)$.

Now, since we are restricting attention to semantics in which every term refers, (i. e. we are not working in free logic,) and since we are now talking about the satisfaction sets for axioms, we can be sure, provided we have not made any errors in our definition of the structures appropriate to the operator in question, that the satisfaction set will never be empty. And now we can make a further stipulation that the function g must be a *choice function* in the sense that $g(M, n, s) \in M$, at least in cases where s is of the appropriate length.

Note that this way out requires that we make a further stipulation, namely that each $\sigma x.\varphi$ occur in exactly one instance of the characteristic axiom schema. Of course, this is merely an inconvenience for our formulation of our axioms, and not a significant restriction.

It's obvious that we can use the same extensionality conditions and still have a dramatic effect on the class of structures we are considering by changing the characteristic axioms for σ. This is what happens, for instance, if we have $(\sigma\alpha)$ as our only extensionality assumption and switch from the ε to the τ calculus, which is a significant change in intuitionistic logic, as we saw in Chapter 11. But we also saw in Chapter 11 that it makes a significant difference if we leave the characteristic axiom fixed and modify the extensionality conditions–for if we are working in the ε-calculus and we make our semantics extensional we end up being able to essentially restrict attention to cases where our lattice of truth-values is a Boolean algebra. This shows that we will quite generally have to pay close attention to both these features. For example, it is common to simply assume as an extra extensionality condition for the description operator ι the rather unnatural extensionality condition

$$\neg\exists!x.\varphi \Rightarrow \iota x.\varphi = (\iota x.x = x)$$

since the characteristic axiom for i is formulated more or less like this:

$$\exists!x.\varphi \Rightarrow (x = \iota x.\varphi \Leftrightarrow \varphi).$$

The justification for adopting this extensionality condition as an axiom is merely formal convenience, but the convenience the defenders often have in mind is the convenience of not needing to work in free logic. The current semantics lets us avoid the unnatural axiom without needing to move to free logic—though admittedly at the price of some other formal inconvenience, namely needing to describe skeleton terms or ground terms.

We now ought to gather all these points together in a more or less formal presentation. We begin with a language $\mathcal{L}$ as described in Chapter 5 for an operator σ. We assume that the behaviour of σ is to be in part determined by a set of *characteristic axioms* for σ, which is such that for every σ-term t there is exactly one characteristic axiom in which t occurs. We adopt the '$Ax(\sigma x.\beta)$' notation as described above.

To get a σ-structure for $\mathcal{L}$, we begin with $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathrm{c}\rangle$, an intuitionistic $\mathcal{L}'$ structure, where $\mathcal{L}'$ is the σ-free sublanguage of $\mathcal{L}$. We say in general that $\mathcal{M}$ is ***σ-attained for $\sigma x.\beta$*** if

$$\{\, d \in D \mid \|Ax(\sigma x.\beta)\|_{\mathcal{M}}^{\varrho(x/d)} = 1 \,\} \neq \emptyset$$

for all ϱ. We say that $\mathcal{M}$ is n *σ-attained* if $\mathcal{M}$ is σ-attained for all σ-terms with n or fewer occurrences of σ. We will, of course, be interested in those $\mathcal{M}$ which are n-σ-attained for every $n \in \omega$. We will call the set $\{ d \in D \mid \|Ax(\sigma x.\beta)\|^{\varrho(x/d)}_{\mathcal{M}} = 1 \}$ the *satisfaction set* for $\sigma x.\beta$, and will denote it by '$ss(\sigma x.\beta)$'.

We also assume that we are given a set of extensionality assumptions for σ, which must include the instances of $(\sigma\alpha)$, but which otherwise might be empty. However, we will insist that the cumulative effect of these axioms must be such that they determine a relation R on the set of σ-terms such that the conditions (1), (2), and (3) are met. First,

(1) If t_1, t_2, and t_3, are all σ-terms, then if Rt_1t_2 and Rt_1t_3, then there is a σ-term t_4 such that Rt_2t_4 and Rt_3t_4.

We inductively define the relation $\approx$ to be the smallest relation such that

$$Rt_1t_2 \implies t_1 \approx t_2$$

$$t_1 \approx t_2 \implies t_2 \approx t_1$$

$$t_1 \approx t_2 \text{ and } t_2 \approx t_3 \implies t_1 \approx t_3.$$

A *normal form t for R* is a term t such that $Rtt_1 \implies t_1 \equiv t$. A σ-term t' *has a normal form t for R* if t is a normal form and $t \approx t'$.

(2) Every σ-term must have a normal form for R.

Now, if we use 'S' to denote the set of σ-terms and 'N' for the set of normal forms for R, conditions (1) and (2) give us an onto map $f : S \to N$. We will also insist that

(3) The map f satisfies Stipulation 1.

If we denote the set of finite sequences of members of D by Seq, then we can define an *interpreting map* E *for* σ to be a map $E : \mathcal{P}(D) \times \mathcal{N} \times Seq \to D$ such that $E(M, n, s) \in M$ when s is a sequence of length m, where the sequence of free occurrences of variables in n is $\vec{x} = \langle x_1, \ldots, x_m \rangle$.

Definition 2. Let E be an interpreting map for σ, and $\langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c} \rangle$ be an intuitionistic structure for $\mathcal{L}'$. If this structure is suitable, we extend the definition of $[t]^{\varrho}_{\mathcal{M}} \in D$ and $\|\varphi\|^{\varrho}_{\mathcal{M}} \in \Omega$ from the terms and formulas of $\mathcal{L}'$ to all of **Terms** and **Wffs** by rewriting Definition 10.1, except that we write $\mathcal{L}$ instead of $\mathcal{L}'$ and we add the clause that if $\mathcal{M}$ is k–σ-attained, then for every σ–term $\sigma x.\beta$ with k or fewer occurrences of s,

$$[\sigma x.\beta]^{\varrho}_{\mathcal{M}} = E(ss(\sigma x.\beta), f(\sigma x.\beta), \langle [t_1]^{\varrho}_{\mathcal{M}}, \ldots, [t_m]^{\varrho}_{\mathcal{M}} \rangle)$$

where $\sigma x.\beta \in (n, \langle t_1, \ldots, t_m \rangle)$. If $\mathcal{M}$ is k–σ-attained for all $k \in \omega$, then $\mathcal{M} = \langle D, \Omega, \mathcal{R}, \mathrm{eq}_{\mathcal{M}}, \mathbf{c}, E \rangle$ is a σ–structure.

If E has the form $E : \mathcal{P}(D) \times \{n\} \times \{*\} \to D$, where n and $*$ are an arbitrary σ-terms and an arbitrary sequence of appropriate length respectively, then $\mathcal{M}$ is a *quasi-extensional* σ–structure. We say that $\mathcal{M}$ is *punctilious* if for every ϱ

$$\|\forall x.(Ax(\sigma x.\varphi) \Leftrightarrow (Ax(\sigma x.\psi))\|^{\varrho}_{\mathcal{M}} \leqslant \mathrm{eq}_{\mathcal{M}}([\sigma x.\varphi]^{\varrho}_{\mathcal{M}}, [\sigma x.\psi]^{\varrho}_{\mathcal{M}}).$$

An *extensional* σ–structure is a punctilious quasi-extensional σ–structure.

It is straightforward to grind through the details of a proof of the following.

Proposition 2. *Every σ-structure is quasi-extensional if and only if the following rule of inference is valid:*

$$(*) \qquad \frac{\forall x.(Ax(\sigma x.\varphi) \Leftrightarrow (Ax(\sigma x.\psi))}{\sigma x.\varphi = \sigma x.\psi}.$$

Every σ–structure is extensional if and only if the following schema is valid:

$$(\sigma\text{ext}) \qquad \forall x.(Ax(\sigma x.\varphi) \Leftrightarrow (Ax(\sigma x.\psi)) \Rightarrow \sigma x.\varphi = \sigma x.\psi. \quad \square$$

It would not be hard to generate more general versions of the Completeness Theorem in Theorem 1 using Proposition 2.

It is also not difficult to describe all the sorts of semantics presented in Chapters 7, 8, 9, and 11 as restrictions of this general scheme in various ways, and we could multiply ad nauseam "results" like the following. The classical cases are those in which we insist that Ω must be a Boolean algebra. (ιext) is equivalent to $\neg\exists! x.\varphi \Rightarrow \iota x.\varphi = (\iota x.x = x)$. And on we could go.

Finally, we will briefly investigate an interpretation we might give to σ which is of interest both because it is an example which makes use of the extra generality of our system, and because it shows that the fact that $\sigma x.(\alpha \wedge \beta)$ and $\sigma x.(\beta \wedge \alpha)$ can get different interpretations is certainly not always a problem. So we don't want to adopt the axioms in Example 1 in all cases. And this calls into question the intuitions which lie behind one of Corcoran and Herring's complaints about Hatcher's original semantics. So let's consider the following operator, which is a slight modification of one invented by Bell as an adaptation to the first order case of a result in type theory due to Andreas Blass.[6] We will call the δ-operator the operator which

[6] The result, which is mentioned in [Bell 1993b], is, in the terminology of that paper, that any localic Hilbertian topos must be Boolean.

has as its characteristic axioms the formulas determined by the following schemas:

If $\varphi \not\equiv \alpha \wedge \beta$

$$\exists x.\varphi \Rightarrow \varphi[x/\delta x.\varphi].$$

If $\varphi \equiv \alpha \wedge \beta$,

$$(\exists x.\beta \Rightarrow \beta[x/\delta x.\varphi]) \wedge (\exists x.(\alpha \wedge \beta) \Rightarrow \alpha[x/\delta x.\varphi]).$$

The δ–operator is obviously a slight modification of the ε—operator. It is of interest because while there are perfectly good non-classical models of the (non-extensional) ε-calculus (even assuming $\vdash c \neq d$), this is not so for δ. We will write $\vdash_\delta$ for derivability in the intuitionistic δ-calculus.

Proposition 3. *Let Σ be a set of sentences such that $\Sigma \vdash c \neq d$ for some constants c and d. Then for any formula γ, $\Sigma \vdash_\delta \neg\gamma \vee \gamma$.*

Proof. Put $\varphi \equiv \alpha \wedge \beta$, where we define $\alpha \equiv x = c$ and $\beta \equiv x = d \vee (x = c \wedge \gamma)$, where x is a variable not free in γ. Obviously, $\vdash \exists x.(\alpha \wedge \beta) \Leftrightarrow \gamma$ and $\vdash \exists x.\beta$. Now our characteristic axiom for $\delta x.\varphi$ tells us that

$$\Sigma \vdash (\exists x.\beta \Rightarrow \beta[x/\delta x.\varphi]) \wedge (\exists x.(\alpha \wedge \beta) \Rightarrow \alpha[x/\delta x.\varphi]),$$

so we have

$$\Sigma \vdash [\delta x.\varphi = d \vee (\delta x.\varphi = c \wedge \gamma)] \wedge (\gamma \Rightarrow \delta x.\varphi = c)$$

and by distributivity

$$\Sigma \vdash (\delta x.\varphi = d \wedge (\gamma \Rightarrow \delta x.\varphi = d)) \vee ((\delta x.\varphi = c \wedge \gamma) \wedge (\gamma \Rightarrow \delta x.\varphi = c))$$

whence $\Sigma \vdash (\gamma \Rightarrow c = d) \vee \gamma$, so that $\neg\gamma \vee \gamma$. $\square$

If it is not too far off the mark to call ε a logical choice function, then it is perhaps not too far off the mark to call δ a logical dependent-choice function. In intuitive terms, to borrow a cute example from John Bell, the δ operator chooses like a kid in an ice cream store. The kid in question might go in with every intention of choosing strawberry ice-cream. But if the store turns out to not have strawberry ice-cream, the kid will take an ice cream anyway—not an arbitrary item from the store, and not, as in what would presumably be the result for the ε-operator in free logic, nothing at all.

At any rate, given the discussion in this chapter, we have a ready-made semantics for this operator. Since our only extensionality assumptions are the $(\delta\alpha)$ axioms, we will want to make the function E depend on precisely the skeleton terms and sequences that were used for the intuitionistic ε- and τ-calculi. We can make the set M be the *satisfaction set* for the term. And we know how to define the structures to ensure that all the characteristic axioms will be satisfied under every valuation in each of them. It would then require only routine modification of the proofs for the intuitionistic ε- and τ-calculi in Chapter 11 to show

Proposition 4. *If we define the δ-structures according to the above suggestion, and write $\models_\delta$ for the usual semantic relations in these structures, we can show that* $\Sigma \vdash_\delta \varphi \iff \Sigma \models_\delta \varphi$. $\square$

Proposition 3 tells us that the Lindenbaum algebra for any theory which includes as a theorem $c \neq d$ is a Boolean algebra.

3. Some Suggestions for Further Research

I would like to close out this thesis by pointing out some of the directions in which the research described in it might be extended.

Some further research might obviously be directed to trying to solve some of the problems left open in earlier chapters, especially in Chapter 11. For example, a precise description of the classes of Heyting algebras which are appropriate to serve as lattices of truth-values for the ε-calculus and the τ-calculus would be interesting.

But the most obvious loose end left by this work is that the last chapter is in a rather undeveloped state. We have seen in Chapters 11 and 12 that adding ε, τ, and δ to the first order intuitionistic predicate calculus has some interesting consequences, and that in particular each of them has consequences we might describe as "changing the logic" of our system. This seems to me to point to two interrelated directions further research might take within the framework described in the present chapter. First, it would be interesting to investigate the question of what other sorts of operators we could invent, and what their consequences would be—in particular, what effect do they have on the logic of the system? An obvious aid to that project, and an interesting project in its own right, is to try to describe (to put the point in a rough and ready way) in more general (i. e. less syntactical) form the reason various changes in our interpreting function E have the effects they do on the lattice of truth-values Ω. This investigation will probably be most effectively carried out in more directly algebraic or category theoretic terms than that of most of the investigation here. For example, we might learn useful facts about the appropriate lattices to be used as lattices of truth-values in the intuitionistic ε calculus by investigating a category whose objects are pairs consisting of a Heyting algebra and a set of subsets of the algebra, each of which has a maximal element.

Finally, it would also be a useful project to *further* generalize the general semantics for term forming operators presented here. My suggestion is that it would be worth investigating the following approach.[7] In Chapter 1 it was briefly mentioned that much of the energy that has been devoted to the semantics of term-forming operators has been focused on trying to generate a semantics for the description operator in free logic because, arguably, free logic is the natural home of that operator. Some authors, Dana Scott in particular, have gone on to suggest ways we might produce a general semantics for term forming operators in classical free logic (cf. [Scott, 1967], [Scott 1970]).

Scott's suggestion, as described in [Scott 1970], is to begin with an indexed set I of possible worlds, and a domain D which includes all the objects in all the possible worlds. If we assume that we have appropriately defined a classical semantics for this so that we simultaneously interpret each formula in all the possible worlds, then, for each φ, $\|\varphi\|^{\varrho}_{\mathcal{M}} : I \to \mathbf{2}$ is a function, so for each formula and each variable we can define a function $d_{\varphi,x} : D \to \mathbf{2}^I$ by putting for each $d \in D$

$$d_{\varphi,x}(d) = \|\varphi\|^{\varrho(x/d)}_{\mathcal{M}} .$$

Since we want to do free logic, we let $[t]^{\varrho}_{\mathcal{M}} : I \to D$ be a *partially defined* function, so there can be possible worlds where, for instance, 'the present king of France' fails

[7] An extremely programmatic discussion of this approach for the classical case is contained in [DeVidi].

to refer. Scott suggests that a term forming operator ought to be interpreted as a map from the set of all the $d_{\varphi,x}$ to the set of interpretations of terms.

Scott's approach already has a number of virtues as it stands, and I don't have much to add to his own defense of it. One feature that is of concern to us, though, is that this approach allows us to express clearly the difference between the extensional and the quasi-extensional versions of the classical ε-calculus. If we write $N\varphi$ for the claim that φ holds in all possible worlds, adopting (Ack) amounts to stipulating that

$$N[(\forall x.\varphi \Leftrightarrow \psi) \Rightarrow \varepsilon x.\varphi = \varepsilon x.\psi],$$

while ε^* calculus is simply one in which

$$N[(\forall x.\varphi \Leftrightarrow \psi)] \implies N[\varepsilon x.\varphi = \varepsilon x.\psi].$$

The second holds as a straightforward consequence of the way the operators are interpreted, while the first requires a stipulation. There is a lesson lurking here about the real content of the stipulation that we will only consider punctilious structures in Chapters 9 and 11.

However, Scott's proposal is obviously not quite what we want. First, we will want to eliminate the restriction to classical logic. And, secondly, Scott's suggestion restricts us to quasi-extensional interpretations of the terms, and we want our interpretations to be flexible enough to handle cases less extensional than that.

Free first order intuitionistic logic is quite a bit more difficult to formulate than free first order classical logic, but an interesting version using an existence predicate is presented by [Scott 1979]. In the same collection which includes that paper

[Fourman and Scott] present a sheaf-theoretic semantics for free intuitionistic logic. It is not hard to make the syntactical modifications necessary to introduce term-forming operators to Scott's system for free intuitionistic logic.[8] It should prove worthwhile to see what results if we can appropriately combine the general semantics for term-forming operators from section two of this chapter with the semantics for free first order intuitionistic logic presented by Fourman and Scott. The result is likely to be complicated and will certainly require application of some rather heavy mathematical machinery, but the result should be a truly general account of term forming operators in first order logic.

[8] I would suggest using Scott's formulation over some others because it requires no restriction on the applicability of Modus Ponens, and so the modifications needed to add term forming operators seem to me more straightforward.

BIBLIOGRAPHY

[Ackermann]

Ackermann, Wilhelm, *Begründung des "tertium non datur" mittels der Hilbertschen Theorie der Widerspruchfreiheit*, Mathematische Annalen **93**, 1–36.

[Asser]

Asser, Günter, *Theorie der logischen Auswahlfunktionen*, Zeitschrift für mathematische Logik und Grundlagen den Mathematik **3** (1957), 30-68.

[Balbes and Dwinger]

Balbes, Raymond and Philip Dwinger, *Distributive Lattices*, University of Missouri Press, Columbia, 1974.

[Barendregt 1981]

Barendregt, H. P., *The Lambda Calculus: Its Syntax and Semantics*, SLFM 103, North-Holland, Amsterdam, 1981.

[Barendregt 1984]

______, *The Lambda Calculus: Its Syntax and Semantics (revised edition)*, SLFM 103, North-Holland, Amsterdam, 1984.

[Barwise, et al.]

Barwise, J., H. J. Keisler, and K. Kunen, eds., *The Kleene Symposium,*, SLFM 101, North-Holland, Amsterdam.

[Bell 1985]

Bell, J. L., *Boolean-Valued Models and Independence Proof is Set Theory (Second Edition)*, Oxford Logic Guides: 12, Clarendon Press, Oxford, 1985.

[Bell 1988]

______, *Toposes and Local Set Theories: An Introduction*, Clarendon Press, Oxford, 1988.

[Bell 1993a]

______, *Hilbert's ε-Operator and Classical Logic*, Journal of Philosophical Logic **22** (1993), 1-18.

[Bell 1993b]

______, *Hilbert's ε-Operator in Intuitionistic Type Theories*, Mathematical Logic Quarterly **39** (1993).

[Bell and Machover]

Bell, J. L. and Moshé Machover, *A Course in Mathematical Logic*, NorthHolland, Amsterdam, 1977.

[Bell and Slomson]

Bell, J. L. and A. B. Slomson, *Models and Ultraproducts: an Introduction*, North Holland, Amsterdam, 1969.

[Benacerraf and Putnam]

Benacerraf, Paul and Hilary Putnam (eds.), *Philosophy of Mathematics: Selected Readings, (second edition)*, Cambridge University Press, Cambridge, 1983.

[Bencivenga et al.]

Bencivenga, Ermanno, Karel Lambert and Bas C. van Fraassen, *Logic, Bivalence and Denotation*, Ridgeview Publishing, Astacadero, Calif., 1986.

[Bernays]

Bernays, Paul, *Appendix to Hilbert's Lecture "The Foundations of Mathematics"*, in [van Heijenoort], pp. 367–392; Translated by Stefan Bauer-Mengelberg and Dagfinn Føllesdal. Originally published 1927.

[Bernays and Fraenkel]

Bernays, Paul and A. A. Fraenkel, *Axiomatic Set Theory*, North-Holland, Amsterdam, 1958.

[Boolos]

Boolos, George, *The Iterative Conception of Set*, Journal of Philosophy **69** (1971). Reprinted in [Benacerraf and Putnam pp. 486–502].

[Bourbaki]

Bourbaki, Nicolas, *Elements of Mathematics: Theory of Sets (second edition)*, Hermann, Paris, 1968 Originally published in French as; *Éléments de Mathématique, Théorie des Ensembles*, Hermann, Paris, 1963.

[Carnap 1956]

Carnap, Rudolf, *Meaning and Necessity: A Study in Semantics and Modal Logic (enlarged edition)*, University of Chicago Press, Chicago, 1956.

[Carnap 1966]

———, *On the Use of Hilbert's ε-Operator in Scientific Theories*, Essays on the Foundations of Mathematics: Dedicated to A. A. Fraenkel on his Seventieth Anniversary (Y. Bar-Hillel, et al., eds.), Magnes Press, Jerusalem, 1966.

[Church]

Church, A., *The Calculi of Lambda Conversion*, Princeton University Press, Princeton, 1941.

[Corcoran and Herring]

Corocoran, John and John Herring, *Notes on a Semantic Analysis of Variable Binding Term Operators*, Logique et Analyse **55** (1971), 644-657.

[Corcoran, Hatcher and Herring]

Corcoran, John, William Hatcher and John Herring, *Variable Binding Term Operators*, Zeitschrift für mathematische Logic und Grundlagen der Mathematik **18** (1972), 177-182.

[Curry and Feys]

Curry, H. B. and R. Feys, *Combinatory Logic*, North Holland, Amsterdam, 1958.

[da Costa 1973]

da Costa, Newton C. A., *Review of* [*Corocoran, Hatcher and Herring*], Zentralblatt für Mathematik **247** (1973), 8–9.

[da Costa 1980]

———, *A Model Theoretic Approach to Variable Binding Term Operators*, Mathematical Logic in Central America (A. Arruda, R. Chaqui, and N. C. A. da Costa, eds.), 1980, pp. 133–162.

[da Costa and Mortensen]

da Costa, Newton C. A. and Chris Mortensen, *Notes on the Theory of Variable Binding Term Operators*, History and Philosophy of Logic **4** (1983), 63–72.

[DeVidi]

DeVidi, David, *Extensionality and term forming operators*, unpublished manuscript.

[Druck and da Costa]

Druck, I. and N. C. A. da Costa, *Sur les "vbtos" selon M. Hatcher*, Comptes rendus de l'Academie des Sciences de Paris **281A** (1975), 741–743.

[Dummett]

Dummett, Michael, *Elements of Intuitionism*, Clarendon Press, Oxford, 1977.

[Eilenberg and Kelly]

Eilenberg, S. and G. M. Kelly, *Closed Categories*, In [Eilenberg et al.], pp. 421–562.

[Eilenberg et al.]

Eilenberg, S., D. K. Harrison, S. Mac Lane and H. Röhrl (eds.), *Proceedings of the Conference on Categorical Algebra, La Jolla 1965*, Springer-Verlag, Berlin, 1966.

[Fine]

Fine, Kit, *Reasoning With Arbitrary Objects*, Aristotelean Society Series, vol. 3, Basil Blackwell, Oxford, 1985.

[Fourman et al.]

Fourman, M. P., C. J. Mulvey, and Dana Scott, eds., *Applications of Sheaves: Proceedings of the Research Symposium on Applications of Sheaf Theory to Logic, Algebra, and Analysis, Durham, July 9-21, 1977*, Lecture Notes in Mathematics, no.753, Springer-Verlag, Berlin, 1979.

[Fourman and Scott]

Fourman, M. P. and Dana Scott, *Sheaves and logic*, In [Fourman et al.], pp. 302-401.

[Frege]

Frege, Gottlob, *Grundgesetze der Arithmetik, vol. 1.* Partial translation in [Furth].

[Furth]

Furth, Montgomery, transl. and ed., *The Basic Laws of Arithmetic. Exposition of The System*, University of California Press, Berkeley, 1964.

[Goos and Hartmanis]

Goos, G. and J. Hartmanis (eds.), *λ-calculus and Computer Science Theory*, Lecture Notes in Computer Science 37, Springer-Verlag, Berlin, 1975.

[Hatcher 1968]

Hatcher, William S., *Foundations of Mathematics*, W. B. Saunders & Co., London, 1968.

[Hatcher 1982]

———, *The Logical Foundations of Mathematics*, Pergamon Press, Oxford, 1982.

[Hazen]

Hazen, Allen, *Natural Deduction and Hilbert's ε-Operator*, Journal of Philosophical Logic **16** (1987), 411–421.

[Hermes]

Hermes, Hans, *Eine Termlogic mit Auswahloperator*, Springer, Berlin, 1965.

[Hilbert 1925]

Hilbert, David, *On the Infinite*, In [van Heijenoort], pp. 367–392; Trans. by S. Bauer-Mengelberg;Originally published as *Über das Unendliche*, Mathematische Annalen **95** (1926).

[Hilbert 1927]

———, *The Foundations of Mathematics*, In [van Heijenoort], pp. 464–479; Trans. by S. Bauer Mengelberg.;From the text of an address delivered in July 1927 at the Hamburg Mathematical Seminar

[Hilbert and Ackermann]

Hilbert, David and William Ackermann, *Grundzüge der theoretischen Logik (second edition)*, Springer, Berlin, 1938.

[Hilbert and Bernays, vol. 1]

Hilbert, David and Paul Bernays, *Grundlagen der Mathematik*, Springer, Berlin, 1934.

[Hilbert and Bernays, vol. 2]

———, *Grundlagen der Mathematik*, Springer, Berlin, 1939.

[Hindley and Seldin]

Hindley, J. Roger and Jonathan P. Seldin, *Introduction to Combinators and λ-calculus.*, London Mathematical Society Student Texts 1, Cambridge University Press, Cambridge, 1986.

[Horn]

Horn, Alfred, *Logic With Truth Values in a Linearly Ordered Heyting Algebra*, Journal of Symbolic Logic **34** (1969), 395–408.

[Kreisel]

Kreisel, Georg, *Hilbert's Programme*, In [Benacerraf and Putnam], pp. 207–248; Originally published Dialectica **12** (1958), 346–372.

[Koymans]

Koymans, C. P. J., *Models of the Lambda Calculus*, CWI Tract 9, Centrum voor Wiskunde en Informatica, Amsterdam, 1984.

[Lambek and Scott]

Lambek, J. and P. J. Scott, *Introduction to Higher Order Categorical Logic*, Cambridge University Press, Cambridge, 1986.

[Lambert 1970]

Lambert, Karel, ed., *Philosophical Problems in Logic: Some Recent Developments*, D. Reidel, Dordrecht, 1970.

[Lambert 1991]

______, ed., *Philosophical Applications of Free Logic*, Oxford University Press, Oxford, 1991.

[Lawvere 1964]

Lawvere, F. W., *An Elementary Theory of the Category of Sets*, Proceedings of the National Academy of Science USA **52**, 1506–1511.

[Lawvere 1972]

Lawvere, F. W., ed., *Toposes, Algebraic Geometry and Logic*, Lecture Notes in Mathematics 274, Springer-Verlag, Berlin, 1972.

[Leisenring]

Leisenring, A. C., *Mathematical Logic and Hilbert's ε-Symbol*, Gordon and Breach Science Publishers, New York, 1969.

[Mac Lane and Moerdijk]

Mac Lane, Saunders and Ieke Moerdijk, *Sheaves in Geometry and Logic*, (Universitext), Springer-Verlag, New York, 1992.

[Maehara]

Maehara, Shôji, *The Predicate Calculus with ε-Symbol*, Journal of the Mathematical Society of Japan, **7** no. 4, (1955), 323–343.

[McLarty]

McLarty, Colin, *Elementary Categories, Elementary Toposes*, Oxford Logic Guides, no. 21, Clarendon Press, Oxford, 1992.

[McKinsey and Tarski 1944]

McKinsey, J. C. C. and Alfred Tarski, *The Algebra of Topology*, Annals of Mathematics **45** (1944), 141–191.

[McKinsey and Tarski 1946]

______, *On Closed Elements in Closure Algebras*, Annals of Mathematics **47** (1946), 122–162.

[Mints 1974]

Mints, G. E., *Heyting Predicate Calculus with Epsilon Symbol*, Selected Papers in Proof Theory, pp. 97–115; Original (in Russian), Zap. Nauch. Sem. Leningrad Otd. Mat. Inst. Akad. Nauk SSSR **40:101** (1974).

[Mints 1990]

———, *Normalization Theorems for the Intuitionistic Systems With Choice Principles*, Mathematical Logic (P. P. Petkov, ed.), Plenum Press, New York, 1990, pp. 59–66.

[Mints (to appear)]

———, *Gentzen-type Systems and Hilbert's Epsilon Substitution Method.*, Logic, Method, and Philosophy of Science (D. Prawitz and D. Westershähl, eds.) (to appear).

[Nerode]

Nerode, A., *Some Lectures on Intuitionistic Logic*, In [Odifreddi], pp. 12–59.

[Odifreddi]

Odifreddi, P., ed., *Logic and Computer Science*, Lecture Notes in Mathematics 1429, Springer-Verlag, Berlin, 1990.

[Padoa]

Padoa, Alessandro, *Logical Introduction to Any Deductive Theory*, In [van Heijenoort pp. 119–123]; transl by J. van Heijenoort from "Essai d'une théorie algébrique des nombres entiers, précédé d'une introduction logique à une théorie déductive quelconque", contributed by Padoa to the Third International Congress of Philosophy, 1900.

[Plotkin 1972]

Plotkin, G. D., *A Set-Theoretic Definition of Application*, School of Artificial Intelligence, Memo MIP-R-95, University of Edinburgh, 1972.

[Plotkin 1974]

———, *The λ-Calculus is ω-Incomplete*, Journal of Symbolic Logic **39** (1974), 313–317.

[Quine]

Quine, W. V., *Set Theory and Its Logic*, Belknap Press of Harvard University Press, Cambridge, 1963.

[Rasiowa]

Rasiowa, Helena, *On ε-Theorems*, Fundamenta Mathematicae **43** (1956), 156–164; *Errata*, Ibid **44** (1957), p. 333.

[Rasiowa and Sikorski]

Rasiowa, Helena and Roman Sikorski, *The Mathematics of Metamathematics, (second edition)*, PWN - Polish Scientific Publishers, Warsaw, 1968.

[Schoenman]

Schoenman, Ralph, ed., *Bertrand Russell: Philosopher of The Century*, George Allen & Unwin, London, 1967.

[Scott 1967]

Scott, Dana, *Existence and Description in Formal Logic*, In [Schoenman], pp. 181-200. reprinted in [Lambert 1991].

[Scott 1970]

———, *Advice on Modal Logic*, In [Lambert 1970], pp. 143–173.

[Scott 1972]

———, *Continuous Lattices*, In [Lawvere 1972], pp. 97–136.

[Scott 1975a]

———, *Combinators and Classes*, In [Goos and Hartmanis], pp. 1–26.

[Scott 1975b]

———, *Some Philosophical Issues Concerning Combinators*, In [Goos and Hartmanis], 1975, pp. 346–366.

[Scott 1976]

———, *Data Types as Lattices*, SIAM Journal of Computing **5** no. 3 (1976), 522–587.

[Scott 1979]

———, *Identity and Existence in Intuitionistic Logic*, In [Fourman et al.], pp. 661–696.

[Scott 1980a]

———, *Lambda Calculus: Some Models, Some Philosophy*, In [Barwise et al.], pp. 223–265.

[Scott 1980b]

———, *Relating Theories of the λ-Calculus*, In [Seldin and Hindley], pp. 403-450.

[Seldin and Hindley]

Seldin, Jonathan P. and J. Roger Hindley, eds., *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, Academic Press, London, 1980.

[van Heijenoort]

van Heijenoort, Jean, ed., *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, Cambridge, 1967.