Logical Foundations
for
Inconsistent Mathematics

Greg O'Keefe

February 2000 - October 2001

I predict a time when there will be mathematical investigations of calculi containing contradictions, and people will actually be proud of having emancipated themselves from consistency.

Ludwig Wittgenstein
Philosophical Remarks

Preface

It would be almost twenty years ago when I decided that I'd like to know about logic. I went to a book shop and bought a book called "Logic" by a guy called Hegel. I couldn't make any sense of it at all, and soon lost interest. I did not know at that stage that *nobody* understands Hegel, and whatever he was talking about in that book, it wasn't Logic!

Much later, wanting to know how computers do what they do, I decided to go to university to study maths, philosophy and computer science. Unfortunately, the computer science people didn't seem very interested in how computers work, just how to write programs, which I already knew. Maths and philosophy were excellent fun though, so I dropped comsci.

There was no logic course when I started at uni. Then Dr. Mark Colyvan arrived and suddenly there were three, *and* one on the philosophy of maths. It was Mark's idea for me to do a thesis on inconsistent maths. Many of the ideas realised in this study were formed under his guidance during the first year of this two year, part-time project.

For my first logic course, the tutor was Dr. J.C. Beall. JC couldn't bear to shelter us from the complicated world of modern Logic. His enthusiastic digressions beyond the elementary level soon had me hooked. I asked JC if he would help me learn Logic. This led to our Saturday sessions at Jane Franklin Hall (where he lived). We fixed his computer problems, swatted for my algebra exams, discussed his research in logical pluralism, ate college food in the sunshine, and read about 50 pages of van Frassen's "Formal Semantics and Logic" [Van71]. Happy days.

JC and Mark also led me to the Automated Reasoning Group's annual Logic Summer School at the ANU. I've now been to this twice, as Dr Jen Davoren invited me to return for a Summer Research Project. While I was there, I had some very helpful discussions on inconsistent maths with Dr John Slaney, and Emeritus Professor Robert K. Meyer.

A co-summer-scholar at ANU was my honours 2000 classmate Paul Hunter. He has made two significant contributions to this project. First, he got deeply into investigating internalised equalities in connection with non-standard analysis.

The result was [Hun00], and hence large parts of chapter 7. Secondly, he gave up much of his mid-semester break reading a draft of this thesis, and emailing his very helpful suggestions.

Mark's idea was to compare Robinson's non-standard analysis with Mortensen's inconsistent analysis. I started by reading about non-standard analysis, then started work on the definitions and results I felt were necessary to present this "properly". In retrospect, I guess it's not surprising that I still haven't acheived it.

The first step in doing mathematics properly, is to get precise about the language you are using. I have been able to do that largely thanks to Dr Peter Trotter's tuition in the course "Languages and Automata". I'm also grateful that he has encouraged me, on several occasions, to be sceptical. I am probably more sceptical than Dr Trotter would consider healthy. I can pick an acceptable proof from a dud, but I sometimes I wonder whether such a proof is really a good reason for believing its conclusion. I also wonder why contradictions shouldn't be true. Hence this study.

Once you have a language, you need to get precise about what the parts of that language *mean*. This tricky business is called Formal Semantics. Back when I started at university, I felt that the maths, philosophy and computer science would all converge to a deep understanding of ... everything. One early sign that this might happen is that the formal semantics I read with JC turns out to be almost the same thing as the universal algebra that I studied under Dr Barry Gardner. Tarski, of course, was a major contributor to both subjects.

Dr Gardner was my co-supervisor initially. Mark generously offered to supervise via email when he was lured to the USA, but my approach was becoming increasingly mathematical. So I imposed on Barry to extend himself beyond his already enormous range of scholarship, into the the dark world of the transconsistent. I owe Barry a huge chunk of my mathematical education. Likewise for Mark and philosophy. Mark also read a late draft at very short notice. If this study is any good, a fair bit of credit must go to these two guys. Of course, where it's not good, no one is to blame but me.

My thanks go to all the people I have mentioned (except for Hegel), as well as my wife Irene and my mother-in-law Bev, who have encouraged me and given me the time to get this done.

Greg O'Keefe
September 2001

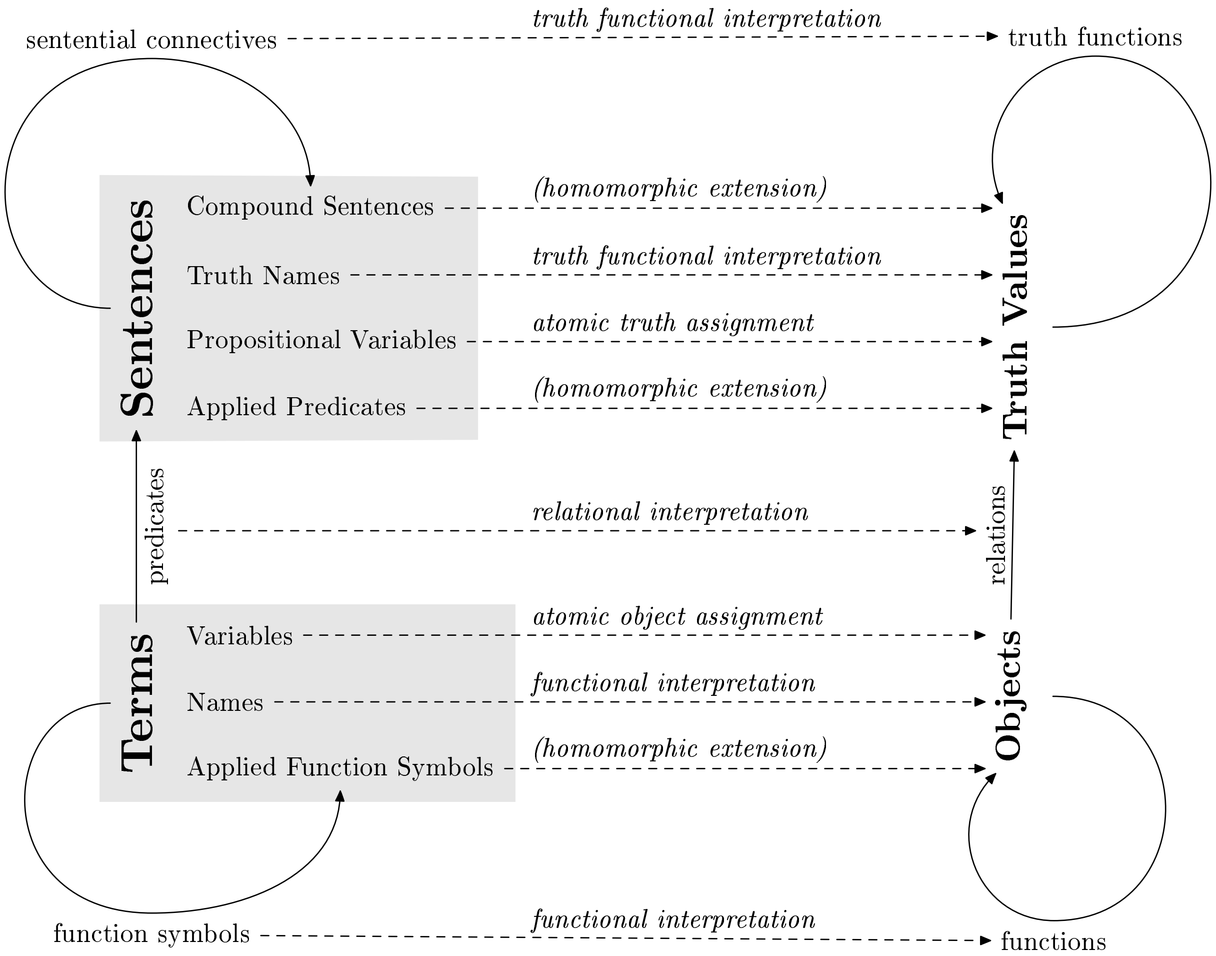

Figure 1: Interpretation of the Language $\mathcal{L}$

Contents

Chapter 1

What?!

The title of this chapter is what I imagine many readers will be thinking when they see the title of this thesis. I hope that by the time we reach Chapter 2, the idea of inconsistent mathematics will make at least a little bit of sense.

This is an exercise in logic. But what is logic? What do logicians do?

The logician seeks to find out what follows from what. We want to be able to extract as much information as possible from any description of facts. From a given set of sentences, we want to find their consequences.

Logic got approximately nowhere until the mid 1800's. Then Boole found that the logical connectives can be treated as operators in an algebra. Logic really got started in 1879, when Frege published *Concept Script*.

Frege really wanted to develop arithmetic from logic alone. For example, with his system of logic, he was able to construct the concept "non-self-identical". Concepts have *extensions*[1], the class of things to which the concept applies. Now everything is identical to itself, so the extension of the concept "non-self-identical" is empty. Frege then defined *zero* to be the class of extensions that can be mapped one-to-one onto this one. Of course he had to define "one-to-one onto mapping", purely in terms of his logic as well. The concept "successor of" (which he also had to define), gave him the natural numbers and hence arithmetic.

Don't be fooled by the brevity of my account. This was a lifetime's work. There can be no handwaving when working on the foundations. Sadly, hardly anyone paid attention to Frege's work while he was alive. Even worse, as the second volume of his *Basic Laws of Arithmetic* was being printed, at his own

[1] Think of these as sets if it helps.

expense, he received a letter from Bertrand Russell showing that the system contained a contradiction. This was the famous Russell paradox: the class of non-self-membered classes is both a member of itself, and not a member of itself.

Russell and Hilbert both attempted programmes similar to Frege's. They also suffered a similar fate[2]: Gödel [Göd31] proved that any consistent formal system that can express arithmetic, is incomplete. There will be sentences that cannot be proved, nor their negations proved, in that system.

Those who wish to do arithmetic[3] in a formal system, face a simple choice then. The system must be inconsistent or incomplete. Until recently, this has been taken to be no choice at all. Mathematicians would not even contemplate working with an inconsistent theory. Why not?

Two distinct and important concepts are routinely conflated by mathematicians: *inconsistency* and *triviality*.

Definition 1.0.1. Let $\mathcal{L}$ be a language with negation $\neg$. A theory Th of $\mathcal{L}$ is *inconsistent* iff there is some formula $A \in \mathcal{L}$ such that $A, \neg A \in Th$. The theory Th is *consistent* otherwise.

The notion of *language* is introduced in Chapter 2, with a definition at 2.2.3. *Theory* is defined in 5.2.9, but roughly, it is a set of sentences that includes all of its consequences.

Definition 1.0.2. A theory $Th \subseteq \mathcal{L}$ is *trivial* iff $Th = \mathcal{L}$.

The following is adapted from [Sto74, Thm. 3.7.1]. Of course, no disrespect is intended. Unlike most authors, Stoll *does* distinguish these two concepts, and considers it worthwhile to prove their equivalence. His proof however rests on assumptions that we consider unjustified.

Flawed Proposition 1.0.3. *Every inconsistent theory is trivial.*

Proofs of this flawed proposition take the following form: Let Th be an inconsistent theory. Then we have a formula A with $A \in Th$ and $\neg A \in Th$. *Every formula is a consequence of* $\{A, \neg A\}$, therefore Th contains every formula.

In short, the proposition is flawed because there are consequence relations such that not every formula is a consequence of $\{A, \neg A\}$.

[2]Russell also had to pay to get *Principia Mathematica* [W&R10] published.

[3]or anything stronger

We will have a lot to say about consequence. In Chapter 3 we define *semantic consequence* (3.7.6). Roughly, this says that the conclusion holds in every case where all the premises do. Chapter 3 (Semantics) is really the core of this study. I take semantic consequence to be prior to deductive consequence. Unlike Hilbert and other "formalists", I prefer to think of mathematical statements as *meaning* something. Once this notion has been made clear, we seek a deductive calculus that gives us the same consequence relation[4].

Chapter 4 briefly surveys deduction, and deductive consequence. A conclusion is a deductive consequence of its premises if it can be *proved* from them. Here, "proof" is a technical term, defined by the particular deductive system in question. This is the key in the move from mathematics to metamathematics. Once you have a definition of proof, you can stop proving things *in* the theory, and start proving things *about* the theory. There are deductive systems where you can not prove an arbitary formula from $\{A, \neg A\}$, although we will not present them here.

In Chapter 5 we explore the general notion of consequence, and its connection with algebraic and topological closures. In Section 6.1 we will give a very simple inconsistent nontrivial theory. The consequence relation under which the theory is closed is a semantic one.

Contradictions do entail everything under most accounts of formal semantics. This is because they allow no interpretation that satisfies a sentence and its negation. If *no* interpretation satisfies $\{A, \neg A\}$, then *every* interpretation that satisfies it satisfies C, for any formula C. Therefore everthing is a consequence of $\{A, \neg A\}$. By employing more than two truth-values, we are able to construct interpretations where contradictions are satisfied, so not everything follows from them.

Why would we want interpretations that satisfy contradictions? Contradictions can't possibly be true, so these interpretations are just depriving us of perfectly good consequences. There are lots of answers to this, but I will give only the ones that seem most important to the present study.

This whole enterprise of formalising mathematical theories, then proving things about them using mathematical techniques, is called *metamathematics.* Usually the game-plan is to prove things using very simple, "undisputable" forms of inference, to show that more controversial properties are

[4]Cynical readers might suggest that I take this position because I did not get enough work done on deduction.

satisfied. For example, Hilbert wanted to show that the “paradise” of infinities discovered by Cantor did not involve any contradictions. Many see these “foundational” exercises as the purest of pure mathematics (and might therefore question their worth). I would argue that this is *applied* mathematics though. It is mathematical logic applied to mathematics.

One way of looking at the present study is as *pure metamathematics.* That is, it doesn't matter whether the consequence relations we construct reflect what *really* follows from what. We just mess around with them because they are interesting. It is comparable to a pure mathematician wondering in times long past, “what if -1 had a square root?” Like that guy (or girl), we just come up with crazy ideas that can't *possibly* have any real life application, and see what we get[5].

There are philosophical logicians though, who argue that contradictions *can* be true. (eg DaCosta [daC74], Priest [Pri87], Rescher [R&B80]) One could see the philosophical logicians as representing the *experimental* branch of logic. (They conduct “thought experiments”) The mathematical logicians persue the *theoretical* aspects of logic. This is where this work fits in. Our role, is to formulate mathematical apparatus that explains the philosophers' “observations”, and enables us to make “predictions” about what follows from what.

The formal semantics that we develop in Chapter 3 is a generalisation of the usual one. As a result we get a definition of *model* (3.7.5) that applies to all the usual mathematical structures and much more besides, including some inconsistent structures. The difference is that our models include a *logic* (defn. 3.5.1). Plenty of people have suggested alternative logics, but they are usually proposing that their logic is the One True Logic. A more recent development is “Logical Pluralism” (see eg [B&R00]), the doctrine that there is no One True Logic, and that different logics are good for different purposes. This is another philosophical motivation for the present work. We claim that rather than simply proving things, the prudent mathematician ought to specify in *which logic* her results hold.

[5]We assume the reader is aware that electrical engineers make constant use of $\sqrt{-1}$.

Chapter 2

Language

Our aim in this chapter is to introduce formal languages. We will also define the formal language that will be used throughout the rest of this work.

The approach taken here is more common in Computer Science and Formal Linguistics literature than it is in logic. I prefer it because it is more concrete and allows us to define languages very concisely using the popular Backus Naur Form (BNF) ([Nau63]).

Much of the material in this chapter comes from [How91].

We will define an object called a *grammar*. The grammar has elements from which we can build sentences, and rules called *productions* determining how the sentences can be built.

The terminology is confusing. Although these "building blocks" are more like English words, we call the set of them the *terminal alphabet*, or simply the alphabet. In some books you also find the sentences of the language called "words".

Grammatical structures like sentences and terms (phrases that name things) are represented by another set of symbols called the *non-terminal alphabet*.

The following example is a quick preview of the detailed presentation that begins in the next section. Consider the sentence "Bill loves Mary." It contains[1] two names and a binary predicate. The grammatical rule that admits this sentence might say something like "The result of placing a binary predicate between two terms is a sentence." In a corresponding formal grammar, we might have non-terminal symbols: σ for sentence; η for names; and ρ for

[1] For simplicity, we neglect the full stop.

binary predicates. The production corresponding to our grammatical rule would then be written

$$\sigma \to \eta\rho\eta$$

In order for "Bill loves Mary" to be a sentence of the language, the formal grammar would also need to contain the productions

$$\rho \to \text{loves}$$

and

$$\eta \to \text{Bill}, \eta \to \text{Mary}$$

This last pair, we would abbreviate by

$$\eta \to \text{Bill} \mid \text{Mary}$$

A *derivation* of the sentence in the grammar is shown by the following notation:

$$\sigma \Rightarrow \eta\rho\eta \Rightarrow \text{Bill } \rho\eta \Rightarrow \text{Bill loves } \eta \Rightarrow \text{Bill loves Mary}$$

Note that this derivation is not unique. It will not bother us here, but it is nice to have a unique derivation for each sentence when working with automata. Roughly speaking, this gives you a *deterministic* automaton.

2.1 Strings

A language over a set of symbols A is a subset of the *strings* over A.

Definition 2.1.1. A *string* over a set A is a finite tuple $\langle a_1, a_2, \ldots, a_n\rangle$ where $n \in \mathbb{N}$ and each $a_i \in A$.

We write A^+ for the set of non-empty strings over A. More precisely, we define

$$A^+ =_{df} \bigcup_{n=1,2\ldots} \{\langle a_1, \ldots, a_n\rangle \ : \ a_i \in A, i = 1 \ldots n\}$$

We will soon find a use for the empty string, so we also define

$$A^* =_{df} A^+ \cup \{\langle\rangle\}$$

We will usually write $a_1a_2\ldots a_n$ for $\langle a_1, a_2, \ldots, a_n\rangle$. Another way of viewing this convention will become possible once we have defined *concatenation*.

Definition 2.1.2. If $b = \langle b_1, b_2, \ldots, b_m \rangle$ and $c = \langle c_1, c_2, \ldots, c_n \rangle$ then we define the *concatenation* of b and c by

$$bc =_{df} \langle b_1, b_2, \ldots, b_m, c_1, c_2, \ldots, c_n \rangle$$

Now concatenation is clearly an associative binary operation on A^+ and A^*, so we have a *semigroup* and a *monoid* respectively. If we neglect the distinction between a and $\langle a \rangle$ for each $a \in A$, then we see that $a_1 a_2 \ldots a_n$ does actually denote $\langle a_1, a_2, \ldots, a_n \rangle$. Also A^+ is the free semigroup generated by A, and A^* the free monoid. See [B&S00, §11].

Notation 2.1.3. We will use the $*$ and $+$ superscripts with arbitary sets of strings, in the obvious way. If B is a set of strings then B^+ will denote the smallest set containing B and closed under concatenation. If we wish to include the empty string, we will write B^*.

2.2 Grammar

Definition 2.2.1. A *grammar* is a tuple $\Gamma = \langle A, N, \Pi, \sigma \rangle$, where

- A is a set of symbols that we call the *terminal alphabet*
- N is a set of symbols, disjoint from A, that we call the *non-terminal alphabet*
- $\Pi \subseteq N^+ \times \{N \cup A\}^*$ is called the *productions*
- $\sigma \in N$ is the *initial element*

We write $u \to v$ to state that the pair $\langle u, v \rangle$ is a member of the productions Π of our grammar. If $u \to v$ and $u \to w$, then we write $u \to v \mid w$.

This notation for the productions has an intuitive reading. We can read

$$\sigma \to P \mid \neg\sigma \mid (\sigma \wedge \sigma)$$

as "A sentence is one of the following: the letter 'P'; a negation sign followed by a sentence; or two sentences with a conjunction sign between them, all enclosed in brackets."

This mechanism for giving languages is known as Backus Naur Form (see [Nau63]). It is equivalent to, yet discovered independently of the phrase structure grammars of Chomsky ([Cho56], [Cho57], [How91]).

We now need an account of exactly how to get a langage from a grammar.

Definition 2.2.2. For $w, w' \in \{A \cup N\}^*$ we write $w \Rightarrow w'$ if there are $x, y \in \{A \cup N\}^*$ and a production $u \to v$ in Π, such that $w = xuy$ and $y = xvy$. We say that w' *derives from* w, and we call the pair $\langle w, w' \rangle$ an *elementary derivation* of Γ.

If $w \Rightarrow w_1 \Rightarrow w_2 \Rightarrow \ldots \Rightarrow w_n = z$ then we also say that z *derives from* w. In this case we write $w \overset{*}{\Rightarrow} z$. We call the n-tuple $\langle w, w_1, \ldots, w_n = z \rangle$ a *derivation* in Γ.

We obtain a language from a grammar by including all the strings of the terminal alphabet that are derived from the initial element.

Definition 2.2.3. The language $\mathcal{L}(\Gamma)$ generated by Γ is

$$\mathcal{L}(\Gamma) =_{df} \{w \in A^* \ : \ \sigma \overset{*}{\Rightarrow} w\}$$

2.3 The Language of Propositional Logics

In this section, I'll define the language of classical propositional logic (CPL) as it is often done in logic books. Then I'll show how this language can be given by a formal grammar as defined above.

The language of CPL is defined as follows. We take a set of propositional variables $P_1, P_2, \ldots$ and the connectives $\neg, \vee, \wedge, \supset, \equiv$. Then the language $\mathcal{S}$ is the smallest set such that

(i) $P_1, P_2, \ldots \in \mathcal{S}$

(ii) if $A, B \in \mathcal{S}$ then $\neg A, (A \vee B), (A \wedge B), (A \supset B), (A \equiv B) \in \mathcal{S}$

Clause (i) gives us the *atomic* sentences. We can include them in the language of a formal grammar by adding the following to the productions

$$\sigma \to P_1 \mid P_2 \mid \ldots$$

The second clause, (ii) can be translated into the following productions

$$\sigma \to \neg\sigma \mid (\sigma \vee \sigma) \mid (\sigma \wedge \sigma) \mid (\sigma \supset \sigma) \mid (\sigma \vee \sigma) \mid (\sigma \equiv \sigma)$$

Our non-terminal alphabet is $N = \{\sigma\}$. Our terminal alphabet is $A = \{(,), \neg, \vee, \wedge, \supset, \equiv, P_1, P_2, \ldots\}$.

Terminology 2.3.1. The language just defined will be refered to as $\mathbb{S}$ throught this study.

The letter $\mathbb{S}$ is chosen to remind us that the language is a *sentential*, not a *predicative* one. That is, the "smallest" building block of the language represents a complete sentence. In the next section we will define a much more useful language, where sentences may be constructed by applying *predicates* to *terms*.

It is common to give the terminal alphabet as several disjoint sets. We may have a set of atomic terms, ie simple names. We may have a set of relation symbols, or perhaps a set for relational symbols of each arity $n \in \mathbb{N}$, and similarly for function symbols. We may also have a separate set of *logical* symbols, whose interpretation (see Chapter 3) is fixed more rigidly than the non-logical ones. Roughly speaking, the logic we are considering at any given moment will be represented by the interpretation of the logical symbols. That will enable us to define validity *under a given logic.* This generalisation of the notion of validity is one of our main objectives.

2.4 The Language of First Order Logics

Here we will define the main language that will be used throughout the remainder of this study. In later chapters, we will sometimes ignore parts of the language, or use more intuition-friendly symbols. By giving a specific concrete language here, I can simplify the presentation of semantics in the next chapter.

The definition looks complicated, but we shall see that it has much in common with ordinary mathematical notation. It might be helpful to look at figure 1 on page (iv), where I have shown pictorially how the parts of this language come together.

Much of the material in this section is drawn from [End72], [Sto74] and [Doe96].

The terminal alphabet A is the union of the following sets:

- propositional variables $\{P_1, P_2, \dots\}$
- sentential connectives
  - of arity 0 (nullary) $\{\top_1, \top_2, \dots, \bot_1, \bot_2, \dots\}$ (truth names)
  - of arity 1 (unary) $\{\neg\}$

- of arity 2 (binary) $\{\vee, \wedge, \supset, \equiv\}$

- logical relation-symbols (logical predicates)
  - of arity 2 (binary) $\{=\}$

- relation-symbols (predicates)
  - of arity 0 (nullary) $\{R_{0,1}, R_{0,2}, \ldots\}$
  - of arity 1 (unary) $\{R_{1,1}, R_{1,2}, \ldots\}$
  - of arity 2 (binary) $\{R_{2,1}, R_{2,2}, \ldots\}$
  - $\vdots$

- variables $\{x_1, x_2, \ldots\}$

- function-symbols
  - of arity 0 $\{n_1, n_2, \ldots\}$ (simple names)
  - of arity 1 (unary) $\{f_{1,1}, f_{1,2}, \ldots\}$
  - of arity 2 (binary) $\{f_{2,1}, f_{2,2}, \ldots\}$
  - $\vdots$

- quantification symbols $\{\forall, \exists\}$

- punctuation[2] : the set containing only the following three symbols (),.

Note that we have three distinct objects here that can stand for a statement. The reasons for this will become clearer in Chapter 3, where we introduce formal semantics. Roughly though, the difference between them is how vulnerable they are to reinterpretation. The nullary sentential connectives, or *truth names* are interpreted by the logic, the nullary relation symbols are interpreted by the model[3] , and the propositional variables are interpreted by a truth-assignment. To be valid in a model, a sentence must be satisfied by all assignments on that model, to be valid in a logic, a sentence must

[2] The usual notation makes it difficult to indicate that the comma symbol is a member of a set. The reader is therefore asked to excuse this rather odd presentation.

[3] Or *model proper*, since later we consider the logic to be part of the model.

satisfy all models on that logic, and to be valid simpliciter, a sentence must be satisfied by all logics[4].

The following productions merely record the roles of the symbols in the formal grammar:

$$\begin{array}{lcll}
\sigma & \to & P_1 \mid P_2 \mid \ldots & \text{propositional variables} \\
\sigma & \to & \top_1 \mid \top_2 \mid \ldots \mid \bot_1 \mid \bot_2 \mid \ldots & \text{truth names (nullary logical connectives)} \\
\rho_0 & \to & R_{0,1} \mid R_{0,2} \mid \ldots & \text{nullary predicates} \\
\rho_1 & \to & R_{1,1} \mid R_{1,2} \mid \ldots & \text{unary predicates} \\
\rho_2 & \to & R_{2,1} \mid R_{2,2} \mid \ldots & \text{binary predicates} \\
\vdots & & & \\
\varepsilon_2 & \to & = & \text{binary logical predicates} \\
\kappa_1 & \to & \neg & \text{unary logical connectives} \\
\kappa_2 & \to & \vee \mid \wedge \mid \supset \mid \equiv & \text{binary logical connectives} \\
\eta & \to & n_1 \mid n_2 \mid \ldots & \text{simple names (nullary function symbols)} \\
\phi_1 & \to & f_{1,1} \mid f_{1,2} \mid \ldots & \text{unary function symbols} \\
\phi_2 & \to & f_{2,1} \mid f_{2,2} \mid \ldots & \text{binary function symbols} \\
\vdots & & & \\
\chi & \to & x_1 \mid x_2 \mid \ldots & \text{(object) variables}
\end{array}$$

We require an existential and a universal quantifier for each of the (object) variables.

$$\omega \quad \to \quad (\exists\chi) \mid (\forall\chi) \quad \text{quantifiers}$$

In addition to the symbols that appear on the left hand sides of these productions, the non-terminal alphabet N also contains τ, which we shall use presently to define the terms of the language.

[4] Note that the "strong" propositional logic K_3, of Kleene has no valid formulae. (see [Pri01, §7.3], [Kle52, §64], [Kle38]) It therefore seems likely that there are no valid formulae in the sense just given.

The remaining productions are

$$\begin{array}{lcll}
\sigma & \to & \kappa_1\sigma & \text{unary compound sentences} \\
\sigma & \to & (\sigma\kappa_2\sigma) & \text{binary compound sentences} \\
\sigma & \to & \rho_0 & \text{nullary predicates} \\
\sigma & \to & \rho_1\tau \mid \rho_2\tau\tau \mid \ldots & \text{applied predicates} \\
\sigma & \to & \tau\varepsilon_2\tau & \text{applied logical predicates (equations)} \\
\sigma & \to & \omega\sigma & \text{quantified sentences} \\
\tau & \to & x_1 \mid x_2 \mid \ldots & \text{(object) variables} \\
\tau & \to & n_1 \mid n_2 \mid \ldots & \text{simple names (nullary function symbols)} \\
\tau & \to & \phi_1(\tau) \mid \phi_2(\tau,\tau) \mid \ldots & \text{applied function symbols}
\end{array}$$

Terminology 2.4.1. The language given by this grammar will be called $\mathfrak{L}$. It will also be useful to have names for two languages defined by the same grammar as this one, except with different initial symols. The language obtained by using ω as the initial symbol will be called the *quantifiers*. The language obtained by using τ as the initial symbol will be called the *terms*.

At times, we will use the formal language in the text, just to show that it can be done. But often we will be sympathetic to the reader by using mnemonic symbols in place of the formal ones, by adding brackets around equations and so on.

Chapter 3

Semantics

In the last chapter, we introduced formal languages. In this chapter, we will show how to attach meaning to the expressions of our languages $\mathcal{S}$ and $\mathcal{L}$.

We are interested in languages because we can use them to *say things*. We care about *statements*, not about questions, exclamations, commands and so on. Statements tell us how things stand in some world. That is they, specify conditions.

The meaning of a statement, for our purposes, is the conditions under which it is true, the conditions under which it is false, and the conditions under which it takes any other mode of veracity. This enables us to give definitions for *consequence*, and *validity*. Although mathematics makes no sense whatsoever without these concepts, we rarely think carefully about them in the way we do about order relations and linear operators and such.

This important notion of consequence may be outlined as follows. Given a collection of statements, and some notion of what conditions are possible in the kind of world we are interested in, we may find other statements that are satisfied by all the worlds of this kind that satisfy our initial collection. These other statements, we call *consequences* of the initial collection.

How are we to go about attaching meaning to the sentences of our languages? The meaning of a sentence depends on its components, and the way these components are put together. Returning to our example from Chapter 2 (page 5), the meaning of "Bill loves Mary" depends on which people the names "Bill" and "Mary" refer to, and what concept the word "loves" denotes. Assume for the moment that the interpretation of these three words is fixed to something like our normal usage. Then although they have the same components, the sentences "Bill loves Mary" and "Mary loves Bill" have the same components but quite distinct meanings. The difference is the way the components are put together.

Our formal languages are given by productions. We will make the meaning of expressions depend on their structure and components by giving interpretation rules that mirror the productions of the languages' grammar.

An interpretation of the sentence "Bill loves Mary" must fix the denotation of the names "Bill " and "Mary", and assign a truth-value to the sentence. This suggests that the meaning of the (two place, infix) predicate "loves" is a map from pairs of objects to truth values.

This is of course the *reference* (Frege's term) or *extension* (Quine) of the relation-symbol. Quine [Qui61] has pointed out that there must be more to meaning than this, otherwise "creature with a heart" and "creature with kidneys" would mean the same thing, as they apply to exactly the same objects. Similarly with Frege's example: "morning star" and "evening star" both denote the planet Venus, but have distinct meanings. In both cases, the two phrases have different *sense* (Frege) or *intension* (Quine) but the same extension or reference. We will confine ourselves to referential meaning here, but sense is not beyond the capabilities of logic. See for example Zalta's [Zal88].

Diagram 1 (page (iv)) summarises the languages and structures (worlds) we will be working with, and the interpretation functions that assign meanings to the expressions of our languages. The remainder of this chapter will be dedicated to explaining and making precise the ideas in this diagram.

3.1 The Universe and the Truth Values

Interpretations of a language are built from two sets: a set of *truth values* and a *universe*.

The truth values are the ways in which a sentence can apply to a world. For example, a sentence might be *true*, *false*, *both true and false*, *known*, *unknowable*, *unknown*, *known to be false*, *proven*, *provable*, and so on. Some of the values in each set of truth values will be *designated*. That is, we consider them ways of being true. The context should make it clear what set of truth values is being used.

Interpretations of the sentential language $\mathcal{S}$ use only a set of truth values. To interpret the predicative language $\mathcal{L}$, we also need a universe. The Universe is the set of things that the language $\mathcal{L}$ is taken to be talking about. The (object) variables range over this set, and the proper names name individuals from it.

Notation 3.1.1. The truth values will be denoted by TV, and the designated truth values by D. We always have $D \subseteq TV$.

The universe will be denoted by U.

Each of the next few sections are dedicated to a kind of expression, and the structure that those expressions are mapped to by the interpretations.

3.2 Sentences and Truth Values

In this section, we will show how give interpretations of the sentential language $\mathcal{S}$ from Section 2.3 (page 8). That is, we will explain the top of diagram 1 (page (iv)) down to "Sentences", "Truth Values" and some of the arrows connecting them. We will call these *truth functional interpretations*. Later, we will extend the truth functional interpretations to interpretations of the full first order language $\mathcal{L}$ we presented in Section 2.4 (page 9).

The sentences of our languages are intended to represent *statements* of natural langages. Natural languages include Russian, Chinese, Hebrew, English, and the symbolically augmented English with which we normally express mathematics. The defining property of statements is that they can be attributed *truth* or $falsity$ (or some other mode of veracity). We will mimic this by mapping the formal sentences into a set of *truth values*, usually including $True$, $False$ and perhaps some other values. The truth or falsity (or whatever) of statements that contain phrases such as "... and ...", "... or ...", "if ... then ..." and "... if and only if ..." depends only on the truth values of the statements that fill the spaces indicated by "...". That is, these phrases are *truth functional.* We accomplish this in our interpretation of the formal language by mapping logical connectives to truth functions. These functions can then be applied to the truth values of the constituents, giving a truth value for the compound sentence.

Definition 3.2.1. A *truth functional interpretation* $\mathcal{M}$ of the language $\mathcal{S}$ (or $\mathcal{L}$) is a triple $\mathcal{M} = \langle TV, D, \ell \rangle$, where ℓ is a function whose domain is the set of sentential connectives of $\mathcal{S}$ and whose range is the functions $TV^n \to TV$ for $n = 0, 1, 2, \ldots$. The arity of $\ell(c)$ is the same as that of c for each logical connective c.

Terminology 3.2.2. Truth functional interpretations are also called *matrices* (hence the $\mathcal{M}$).

A truth functional interpretation of $\mathcal{S}$ will also be called a *logical interpretation* of $\mathcal{S}$, an *interpretation* of $\mathcal{S}$, or a *propositional logic*.

They are lucky structures indeed to have so many names. As we shall see in Section 3.5, a logical interpretation of $\mathcal{L}$ involves more than a truth functional interpretation, and (Section 3.7.5) an interpretation of $\mathcal{L}$ involves more than a logical interpretation of $\mathcal{L}$.

Notation 3.2.3. We will write $\neg^{\mathcal{M}}$ for the truth function $\ell(\neg)$ corresponding to the logical connective $\neg$, $\vee^{\mathcal{M}}$ for the truth function $\ell(\vee)$ corresponding to the logical connective $\vee$, and so on.

We call a function from the propositional variables to the truth values an *atomic valuation*. Intuitively, we can see this as a specification of a world. Wittgenstein [Wit22] and Russell [Rus18, Rus24] developed a philosophical theory called *Logical Atomism*. In brief, this theory says that the world consists of a collection of logically independent facts: "The world is everything that is the case" [Wit22, §1]. An atomic truth assignment then, tells you what is the case. The truth values of the compound sentences depend on the truth values of the atomic sentences.

Given a truth functional interpretation, we can uniquely[1] extend an atomic valuation to all the compound sentences of the language formed from propositional variables and logical connectives. We will call such an extention a *valuation* of $\mathcal{S}$.

Definition 3.2.4. Let $\mathcal{M}$ be a truth functional interpretation of $\mathcal{S}$, and v be a function $v : \{P_1, P_2, \dots\} \to TV$. We call v an *atomic valuation*. Let $\overline{v}$ be given by:

$$\begin{array}{rcl} \overline{v}(P_i) & = & v(P_i) \\ \overline{v}(\neg\varphi) & = & \neg^{\mathcal{M}}\overline{v}(\varphi) \\ \overline{v}(\varphi \vee \psi) & = & \overline{v}(\varphi) \vee^{\mathcal{M}} \overline{v}(\psi) \\ \vdots & \vdots & \vdots \end{array}$$

for each sentence $\varphi, \psi \in \mathcal{S}$. We call $\overline{v}$ an $\mathcal{M}$-*valuation* of $\mathcal{S}$.

[1] see [B&S00] lemma 10.6

Terminology 3.2.5. A valuation of $\mathcal{S}$ will also be called a *truth assignment* of $\mathcal{S}$, as it gives a truth value to every sentence of $\mathcal{S}$. It will also be called an *admissible* valuation, since they are all admissible for our purposes, in a sense to be defined later (3.5.1). However a valuation of $\mathcal{L}$ is a more complicated beast, that we will get to in the next section.

Notice that this makes the set of truth assignments exactly the set of homomorphisms from the language, viewed as a term algebra, to the matrix.

We are now ready to introduce the central concept of logic: *consequence*. This definition is of a special kind of consequence, *semantic consequence*, and only applies to the propositional language $\mathcal{S}$. For semantic consequence over $\mathcal{L}$ we have more work to do, but we will get there in definition 3.7.6. For an exploration of consequence in its full generality, you have to wait until Chapter 5.

Definition 3.2.6. Let L be a propositional logic, $\overline{v}$ an L-valuation of $\mathcal{S}$. We say that a sentence $\varphi \in \mathcal{S}$ is *satisfied* by $\overline{v}$ iff $\overline{v}(\varphi) \in D$. The sentence φ is a *semantic consequence under the logic L* of a set of sentences $\Gamma \subseteq \mathcal{S}$ iff every L-valuation that satisfies each member of Γ also satisfies φ. If a sentence φ is a semantic consequence of the empty set, we say that φ is *valid* under the propositional logic L.

Notation 3.2.7. We will use the notation $\Gamma \models_L \varphi$ to say that $\varphi \in \mathcal{S}$ is a semantic consequence of $\Gamma \subseteq \mathcal{S}$, under the propositional logic L. To say that φ is valid under this logic, we write $\models_L \varphi$. We will omit the subscript whenever we can get away with it.

Terminology 3.2.8. If a sentence is valid under a logic, we will sometimes say that it *holds* under that logic. Otherwise, we may say that it *fails* under that logic.

We now have a set of valuations for each propositional logic. We would like a definition of semantic consequence for the full first order language $\mathcal{L}$, so we need a definition of logical interpretation for that language, and for each logical interpretation, a set of valuations.

Notice how the recursive definition of the language parallels the recursive definition of the interpretation. Consider a sentence $\neg\varphi$. The "outermost" production in its derivation is $\sigma \to \neg\sigma$. Its interpretation is defined by $\overline{v}(\neg\varphi) = \neg^{\mathcal{M}}\overline{v}(\varphi)$. The terminal symbol introduced on the right hand side of the production ($\neg$) is mapped to a function ($\neg^{\mathcal{M}}$), which is applied to the interpretation of the recursively defined component (φ). This corresponds to the non-terminal (σ) on the right hand side of the production[2]. This pattern will continue as we extend the interpretations to $\mathcal{L}$.

3.3 Terms and Objects

Consider just the terms of the language $\mathcal{L}$, depicted in the bottom left of the diagram on page (iv). We can obtain this language by substituting τ for σ as the initial symbol of the grammar. This language is very similar to the sentential language $\mathcal{S}$ we considered in the previous section. Where the truth functional interpretation assigns truth functions to logical connectives, this interpretation will assign functions to function-symbols.

Definition 3.3.1. A *functional interpretation* $\mathcal{F}$ of the language $\mathcal{L}$ is a pair $\langle U, \iota \rangle$, where U is a universe, and ι is a function. The domain of ι is the set of function-symbols of $\mathcal{L}$, its range is the functions $U^n \to U$ for $n = 0, 1, 2, \ldots$. The arity of $\iota(f)$ is the same as that of f for each function symbol f.

Notation 3.3.2. We will write $f^{\mathcal{F}}$ for the function $\iota(f)$ denoted by f under the functional interpretation $\mathcal{F}$.

When we were interpreting $\mathcal{S}$ we had valuations taking sentences to truth values. Here we need to take terms to objects. The approach will be very slightly different though, in order to improve our account of quantification (Section 3.6).

Definition 3.3.3. Let $\mathcal{F}$ be a functional interpretation of $\mathcal{L}$, and $\underline{a}$ be $\underline{a} = \langle a_1, a_2, \ldots \rangle$, with each $a_i \in U$. That is $\underline{a} \in U^{\mathbb{N}}$. We call $\underline{a}$ a *sequence of objects*. Let the function $\overline{a}$ from the terms into the universe, be given by:

$$\begin{aligned} \overline{a}(x_i) &= a_i \\ \overline{a}(n_i) &= n_i^{\mathcal{F}} \\ \overline{a}(f_{m,n}(\tau_1, \ldots, \tau_m)) &= f_{m,n}^{\mathcal{F}}(\overline{a}(\tau_1), \ldots, \overline{a}(\tau_m)) \end{aligned}$$

for each $i, m, n \in \mathbb{N}$. We call $\overline{a}$ an *object assignment* of $\mathcal{L}$.

Terminology 3.3.4. Sometimes we will think of $\underline{a}$ as a function $\underline{a} : X \to U$ given, as above, by $\underline{a}(x_i) = a_i$. In these cases we will call $\underline{a}$ an *atomic object assignment*.

[2] I'd very much like to be able to provide an animation of this. It is a simple idea, but difficult to express.

Why am I making this rather fine distinction? The usual account of quantification makes it quite clear how sentences containing quantifiers should be evaluated, but they do not give any object or function to which each quantifier corresponds. You could do this with atomic object assignments as functions. This does not seem acceptable to me. I want something outside the language, rather than a map from part of the language into the universe. This is so that I can construct a denotation for each quantifier, that is completely outside the language, as good dentotations should be.

Notation 3.3.5. We will align our notation with the (weird) model-theoretic norm by writing $\varphi[\underline{a}]$ for $\overline{a}(\varphi)$.

Again, we have the set of homomorphisms from one algebra into another. The language is again a term algebra, and the universe with the functional interpretation is an algebra of the same type.

3.4 Predicates and Relations

Relations are usually thought of as subsets of Cartesian products. So if R is an $n-$ary relation on a set S, then $R \subseteq S^n$. When n is 2, we say that $a \in S$ is R-related to $b \in S$ iff $\langle a, b \rangle \in R$.

This, we claim, is a kludge that hinders full generality in the study of logic. From the denotations of the name symbols b and m, and the denotation of the relation-symbol L, one ought to be able to determine the truth value of the sentence Lbm. The method just outlined will only work if the set of truth values has two or less members. Hence, the objects that will be denoted by our relation-symbols are defined as follows.

Definition 3.4.1. An n-ary *relation* R on the set U for the truth values TV is a function $R : U^n \to TV$.

We will occasionally use relations in the usual sense when working *on* the formal semantics, but the relations *in* our models are always of the kind just defined. We have no need for relations between distinct sets, but it would be straight-forward to define them, should the need arise.

The designated truth values can be seen as the ways in which a relation can hold, the non-designated truth values, ways in which a relation can fail to hold.

Definition 3.4.2. A *relational interpretation* $\mathcal{R}$ of the language $\mathcal{L}$ is a triple $\langle U, TV, \jmath\rangle$, where U is a universe, TV a set of truth values, and $\jmath$ is a function. The domain of $\jmath$ is the predicates of $\mathcal{L}$, its range is the functions $U^n \to TV$ for $n = 0, 1, 2, \ldots$. The arity of $\jmath(R)$ is the same as that of R for each relation symbol R.

Notation 3.4.3. We will write $R^{\mathcal{R}}$ for the relation $\jmath(R)$ denoted by R under the relational interpretation $\mathcal{R}$.

Note that our language $\mathcal{L}$ has no relational variables. These could easily be added to the language, along with relational quantifiers. Relational assignments would be required to interpret this extended language. What we would then have is "second order logic".

3.5 Logical Predicates and Possibility

The interpretations that we are constructing can plausibly be seen as *possible worlds*. I will not try to make this idea exact, because it is difficult and controversial. See, for example [Kri80].

This leads us to a different intuitive idea of consequence: φ is a consequence of Γ if it is impossible for φ to fail in a world where all the sentences in Γ hold.

Now, is it possible for $a \neq a$? In most contexts, no, its not possible, though in Section 6.1 we will show this to be a valid formula of a non-trivial theory. But if '=' can be mapped to *any* binary relation, we get (preempting definition 3.7.1 of valuations of $\mathcal{L}$) "impossible" valuations like the following.

Let $U = \{a\}$, $TV = \{True, False\}$, $D = \{True\}$, $\neg^{\mathcal{M}}(True) = False$ and $\neg^{\mathcal{M}}(False) = True$ and $n_0^{\mathcal{F}} = a$. Nothing shocking so far. But we can let $=^{\mathcal{R}} (a, a) = False$, and hence we have

$$\begin{aligned}\overline{va}(\neg(n_0 = n_0)) &= \neg^{\mathcal{M}}(=^{\mathcal{R}} (n_0^{\mathcal{F}}, n_0^{\mathcal{F}})) \\ &= \neg^{\mathcal{M}}(=^{\mathcal{R}} (a, a)) \\ &= \neg^{\mathcal{M}}(False) \\ &= True\end{aligned}$$

What can we do to make our semantics say that $a \neq a$ is impossible? We can't attach the $=$ sign to a particular relation, because we have no fixed universe on which to define a relation.

We will simply *exclude* these impossible valuations.

Definition 3.5.1. A *first order logic* is a truth functional interpretation $\mathcal{M}$, with a quantitative interpretation $\mathcal{Q}$ (defn. 3.6.2) and an informally specified list of *possibility constraints* that define a subset of the valuations of $\mathcal{L}$ into TV. This subset is called the *admissible* valuations.

This general approach to semantics for logical predicates is due to Varzi [Var01], who also uses it to solve the problem of the following notorious argument:

> This apple is red, therefore this apple is coloured.

If we translate this into a formal language in the usual way, we get a premise Ra (this apple is red) and a conclusion Ca (this apple is coloured). There are interpretations where R denotes a function that takes the denotation of a to a designated truth value, yet C's denotation takes a's to a non-designated value. For example (using the usual two valued logic, and using the same name for the apple in both the object language and the meta-language):

$$\overline{va}(Ra) = R^{\mathcal{R}}(a) = True$$

but

$$\overline{va}(Ca) = C^{\mathcal{R}}(a) = False$$

The logic therefore counts the argument invalid. But surely it's impossible for something red not to be coloured?

We may rectify this by making R and C logical predicates, and imposing the constraint that the relation denoted by C must take to designated truth values, all those objects taken to designated truth values by the relation denoted by R. Varzi calls[3] this a "colour logic".

To put this symbolically, the relational interpretation $\mathcal{R}$ is admissible iff

$$\{x : R^{\mathcal{R}}(x) \in D\} \subseteq \{x : C^{\mathcal{R}} \in D\}$$

Rejecting inadmissible relational interpretations in turn excludes certain (impossible) valuations.

[3]Though not very seriously I think.

Although in this study we are content to give possiblility constraints informally, it may be fruitful to explore ways, such as equations in universal algebra, of giving them more rigorously.

The sentential connectives could also be treated this way. For example, we may just specify that $P \wedge Q$ is designated whenever P and Q are, rather than giving a specific function $\wedge^{\mathcal{M}}$.

3.6 Quantification

The following account of quantification is a bit novel. If there is something badly wrong with it, the blame rests with the author, not his supervisor, nor the publications cited at the beginning of this chapter.

Recall that in Section 3.3.4 we decided that an atomic object assignment, and a sequences of objects are pretty much the same thing. Consider a formula Rx_kx_m where R is a binary predicate symbol. We will obtain a truth value for the formula by

(i) applying a sequence of objects $\underline{a}$ to x_k and x_m to get a pair of objects, a_k and a_m,

(ii) applying a relational interpretation to R to get a relation, $R^{\mathcal{R}}$,

(iii) then applying $R^{\mathcal{R}}$ to a_k and a_m to yield a truth value.

More compactly, we have

$$R^{\mathcal{R}}x_kx_m[\underline{a}] = T_1 \in TV$$

Given a truth functional interpretation, we could also obtain a truth value for a compound formula such as $Rx_kx_m \vee Sx_m$, where S is a unary predicate symbol. Similarly, formulae involving function symbols can be evaluated if we have a functional interpretation.

Holding the relational, functional and truth functional interpretations fixed, we could take another sequence of objects and obtain another truth valuation for the formula

$$R^{\mathcal{R}}x_kx_m[\underline{b}] = T_2 \in TV$$

perhaps with $T_2 \neq T_1$.

In other words, we can see an open[4] formula as denoting a function that takes a sequence of objects to a truth value. Symbolically

$$R^{\mathcal{R}} x_k x_m[\underline{a}] = f(\underline{a})$$

where f is a function $f : U^{\mathbb{N}} \to TV$. This makes the model-theorists' notation (see 3.6.4) look better motivated. Maybe it is not a good idea to think about the denotations of partially interpreted formulae. If it doesn't disturb you unduly, we can express the idea that the formula denotes a function (f) by saying that

$$R^{\mathcal{R}} x_k x_m \in TV^{U^{\mathbb{N}}}$$

Closed formulae need not be excluded from this treatment; they simply denote constant functions of this type. No matter what sequence of objects you apply them to, you get the same truth value in return. We will specify this in the possibility constraints of most of our logics.

How then are we to interpret the formula $(\forall x_k)Rx_kx_m$? This is itself an open formula, since x_m occurs free. The quantifier $(\forall x_k)$ ought to denote a function that takes the denotation of Rx_kx_m to the denotation of $(\forall x_k)Rx_kx_m$. (Recall our little rant about this at the end of Section 3.2, page 17.) That is, $(\forall x_k)$ denotes something that takes one function from sequences of objects to truth values, to another function from sequences of objects to truth values.

Terminology 3.6.1. Wherever in this chapter we have talked about truth values, now think of these as constant functions from sequences of objects to truth values.

Definition 3.6.2. A *quantitative interpretation* $\mathcal{Q}$ of the language $\mathcal{L}$ is a triple $\langle U, TV, \sharp\rangle$, where U is the universe, TV the truth values, and $\sharp$ is a function. The domain of $\sharp$ is the set of quantifiers of $\mathcal{L}$, its range is the functions $TV^{U^{\mathbb{N}}} \to TV^{U^{\mathbb{N}}}$.

Notation 3.6.3. We will write $(\forall x)^{\mathcal{Q}}$ for the function $\sharp((\forall x))$ denoted by $(\forall x)$ under the quantitative interpretation $\mathcal{Q}$.

But not every quantitative interpretation will do. We want the quantifier $(\exists x_k)$ to *mean* "there is some object, call it x_k, such that ..." and similarly

[4] I owe the reader several definitions here, but I hope these notes are enough. A formula is *open* if it contains free variables. A variable in a formula is *free* if it is not bound by a quantifier. For example x is bound in $(\forall x)(Py \wedge Px)$, but y is free.

for the universal quantifiers. We need some possibility constraints to exclude inapropriate interpretations of the quantifiers.

The following definition will help us to formulate these constraints, but first, a bit more notation.

Notation 3.6.4. We will write $\underline{a}_b^k$ for $\langle a_1, \dots, a_{k-1}, b, a_{k+1}, \dots \rangle$, the sequence that is the same as $\underline{a}$, except with b in position k.

Again, this is fairly close to what model-theorists usually do.

Definition 3.6.5. Let $f : U^{\mathbb{N}} \to TV$ be a function. If for each $\underline{a} \in U^{\mathbb{N}}$, and each $b \in U$

$$f(\underline{a}) = f(\underline{a}_b^k)$$

then we say that f is *fixed in the k-th place.*

If f is fixed in the k-th place for each $k \in K \subseteq \mathbb{N}$ we say f is *fixed in the places K*, and *fixed everywhere but* $\mathbb{N} \setminus K$.

Here then is a proposed list of possibility constraints that should give a fairly sensible interpretation of the quantifiers in a first order logic.

Requirements 3.6.6. The truth values TV with $\wedge^{\mathcal{M}}$ and $\vee^{\mathcal{M}}$ should form a lattice, with partial order $\preceq$, and for each formula $\varphi \in \mathcal{L}$ the function $f : U^{\mathbb{N}} \to TV$ denoted by φ should satisfy the following

(i) f is fixed in the k-th place unless the variable x_k is in φ,

(ii) $(\forall x_k)^{\mathcal{Q}}(f)$ and $(\exists x_k)^{\mathcal{Q}}(f)$ are fixed in the k-th place

(iii) if $g = (\exists x_k)^{\mathcal{Q}}(f)$ then there is some *representative* $e \in U$ such that for all $\underline{c} \in U^{\mathbb{N}}$,

$$g(\underline{c}) = f(\underline{c}_e^k)$$

where e is maximal, in a sense that I shall discuss presently,

(iv) and similarly for the universal quantifiers, with a minimal representative a.

Intuitively, we want e to be the element that makes the unquantified formula φ "as true as possible", and a the one that makes it "as false as possible".

This is not difficult if φ has one or less free variables, and the quantified formula has none. In that case, φ effectively denotes a unary relation (predicate) $P : U \to TV$. Since the truth values form a lattice, we can define an ordering of U by

$$b \leq_P c \text{ iff } P(b) \leq P(c)$$

and require that the representative by minimal/maximal under this order.

What happens when (say) $(\exists x_k)\varphi$ has free variables? Its probably best to try to order the sequences of objects, then choose the representative from a maximal sequence. Each sequence takes the denotation of any formula to a truth value (constant truth function). So the sequences can be ordered in much the same way.

Finding an exact sense in which the k-th element of a maximal sequence, is itself maximal is left as an exercise for the reader (*hint*: Are projections order homomorphisms?).

Your second exercise is the following:

Wish 3.6.7. *We would like a proof that the logic with the classical truth functional interpretation and the possibility constraints 3.6.6, or something like them, is classical first order logic.*

3.7 Truth and Consequence

We are now ready to define *truth assignment* (valuation) for the full first order language $\mathcal{L}$. This will give us notions of truth and consequence for that language.

Definition 3.7.1. Let

- $\mathcal{L}$ be our first order language,
- $\mathcal{M}$ a truth functional interpretation of $\mathcal{L}$ into TV
- v an atomic truth assignment of $\mathcal{L}$
  (in the same sense as those for $\mathcal{S}$, see 3.2.4, 3.2.5)
- $\mathcal{F}$ a functional interpretation of $\mathcal{L}$
- $\underline{a}$ a sequence of objects $\in U^{\mathbb{N}}$
- $\mathcal{R}$ a relational interpretation of $\mathcal{L}$

- $\mathcal{Q}$ a quantitative interpretation of $\mathcal{L}$

where the universes U and sets of truth values TV for all these things are the same.

Let $\overline{va}$ be given by:

$$\begin{aligned}
\overline{va}(P_j) &= v(P_j)\\
\overline{va}(\top_j) &= \top_j^{\mathcal{M}}\\
\overline{va}(\bot_j) &= \bot_j^{\mathcal{M}}\\
\overline{va}(\neg\varphi) &= \neg^{\mathcal{M}}\overline{va}(\varphi)\\
\overline{va}(\varphi \vee \psi) &= \overline{va}(\varphi) \vee^{\mathcal{M}} \overline{va}(\psi)\\
\vdots \quad &\vdots \quad \vdots\\
\overline{va}(x_i) &= a_i\\
\overline{va}(n_i) &= n_i^{\mathcal{F}}\\
\overline{va}(f_{i,j}\alpha_1, \dots, \alpha_i) &= f_{i,j}^{\mathcal{F}}(\overline{va}(\alpha_1), \dots, \overline{va}(\alpha_i))\\
\overline{va}(R_{i,j}\alpha_1, \dots, \alpha_i) &= R_{i,j}^{\mathcal{R}}(\overline{va}(\alpha_1), \dots, \overline{va}(\alpha_i))\\
\overline{va}(\alpha_1 = \alpha_2) &= \overline{va}(\alpha_1) =^{\mathcal{R}} \overline{va}(\alpha_2)\\
\overline{va}((\forall \upsilon)\varphi) &= (\forall \upsilon)^{\mathcal{Q}}(\overline{va}(\varphi))
\end{aligned}$$

for each $i, j \in \mathbb{N}$, $\varphi, \psi \in \mathcal{L}$, each variable υ, and all terms $\alpha_1, \dots, \alpha_i$. We call $\overline{va}$ a *valuation* of $\mathcal{L}$.

The notation $\overline{va}$ is intended to convey the idea that we are combining the truth valuation v and the sequence of objects $\underline{a}$, and extending this combiniation using all the interpretation functions. Happily, it also reminds us that it is a $\overline{va}$luation.

Worry 3.7.2. I am equivocating about whether the atomic object assignment $\overline{a}$ is applied to the variables before or after all the interpretations are applied to their respective components. If $\overline{a}$ is applied last, that implies that the truth functions are defined not over truth values, but over functions into the truth values. I don't *think* we can get into trouble because of this, but a proof, or even a clear formulation of the problem would be reassuring.

Definition 3.7.3. A valuation that meets the possibility constraints of its logical interpretation is called an *admissible* valuation of that logic.

Definition 3.7.4. We say that a valuation $\overline{va}$ *satisfies* a sentence $\varphi \in \mathfrak{L}$ iff $\overline{va}(\varphi) \in D$.

Definition 3.7.5. A first order logic L with a non-logical interpretation over the truth values of L will be called an *interpretation* or a *model.* If every admissible valuation of a model $\mathfrak{M}$ satisfies a set of sentences $\Gamma \subseteq \mathfrak{L}$ we say that $\mathfrak{M}$ is a *model of* Γ, And we write $\mathfrak{M} \models \Gamma$.

Note that in the notation $\mathfrak{M} \models \Gamma$, we are not omitting any subscript. Everything that matters is packed into $\mathfrak{M}$. In more "conventional" accounts of models in non-classical logics, $\mathfrak{M}$ would stand only for the non-logical interpretation, and the logic L would be indicated by a subscript $\mathfrak{M} \models_L \Gamma$. This doesn't make any difference most of the time, but Chapter 7 is an extended example of a structure that is a model in our sense, but not in the more usual sense.

At last we reach the goal of this chapter, to define semantic consequence for our first order language $\mathfrak{L}$.

Definition 3.7.6. The sentence φ is a *semantic consequence* (under the first order logic L) of a set of sentences $\Gamma \subseteq \mathfrak{L}$ iff every admissible valuation of L that satisfies each member of Γ, also satisfies φ. If a sentence φ is a semantic consequence of the empty set, we say that φ is *valid* (under the logic L).

Notation 3.7.7. We will use the notation $\Gamma \models_L \varphi$ to say that $\varphi \in \mathfrak{L}$ is a semantic consequence of $\Gamma \subseteq \mathfrak{L}$, under the first order logic L. If $\Gamma = \{P_1, P_2, P_3\}$ we sometimes write $P_1, P_2, P_3 \models \varphi$ for $\Gamma \models \varphi$. To say that φ is valid under this logic, we write $\models_L \varphi$. As usual, when no clarity is lost the subscript goes.

Chapter 4

Deduction

This chapter very briefly introduces deduction.

Many people would define Logic as the study of deduction. But deduction in our opinion, is only part of what Logic is about. Here, we take a much broader view of Logic that also includes the theory of language, formal semantics, the theory of computation and parts of Philosophy. Indeed we see deduction as secondary to formal semantics.

Semantics tells us what *really* follows from what. Once we have a semantic consequence relation, we then set out to find a formal deductive apparatus that yields the same consequence relation.

Another way of expressing this view is to say that deductive systems are proposed solutions to the following *decision problem*

instance: A subset Γ and a member C of $\mathfrak{L}$, and a first order logic[1] L

question: Is C a semantic consequence of Γ?

The problem is undecidable for $\mathfrak{L}$ and the usual consequence relation [Jef91, Chapter 8], so we have to make do with partial solutions.

The answer (yes or no) is not all that matters though. We also want the workings of this machine to be *convincing*. Ideally, given an instance of the problem, the machine will return a construction that we can plausibly think of as a *proof*[2].

---

[1] This really prevents the problem from being a "problem" in the technical sense, because the logic includes informally specified possibility constraints. It is not clear that these can captured by an encoding scheme.

[2] This requirement can equally be extended to other "problems" in automata theory. A

To give a system of deduction then, is to *define proof.* This is the final step in formalising a mathematical subject. It enables us to "step back" from that subject, and treat it as a mathematical object. Hence, if everyone agrees on the formalisation, no dispute is possible about the theorems of that subject. Nothing depends on intuition. Abstract, infinite and disputable ideas are "reduced" to concrete, finite and undisputable formal derivations.

4.1 Kinds of Deductive System

There are four widely used types of deductive system: Hilbert Systems; Natural Deduction; Sequent Calculi; and Semantic Tableaux.

The most famous example of a Hilbert style deductive system is that of Russell and Whitehead's *Principia Mathematica* [W&R10].

The following definition is from [End72, §2.4, p. 103] ($\Lambda \subseteq \mathfrak{L}$ is the set of *axioms*)

Definition 4.1.1. A *deduction of* φ *from* Γ is a sequence $\langle \alpha_0, \ldots \alpha_n \rangle$ of formulae such that $\alpha_n = \varphi$ and for each $i \leq n$ either

(i) α_i is in $\Gamma \cup \Lambda$, or

(ii) for some j and k less than i, α_i is obtained by modus ponens from α_j and α_k (ie, $\alpha_k = \alpha_j \supset \alpha_i$)

A distinctive feature of these systems is that there is only one rule of inference, modus ponens (from A and $A \supset B$ deduce B). This is helpful when proving properties such as *soundness* and *completeness*, because you only need to prove that the property is preserved by that one rule.

Natural deduction proofs have some resemblance to the informal proofs you find on blackboards and in maths texts. A proof is a sequence of formulae, as for Hilbert systems, but there are many more rules of inference.

For each of the logical symbols, there is an introduction and an elimination rule. For example, an introduction rule for $\wedge$ might (ought to) say, that if a proof contains A and also contains B, then $A \wedge B$ may be added to the proof.

---

machine that returns a Hamiltonian circuit on a graph you have given it is a more useful machine than one that simply tells you that such a circuit exists. Generally, the algorithm for solving a problem ought to help to *explain* the problem. Unfortunately "explanation" is even more difficult to define than "problem".

The elimination rule allows you to add A or B if you already have $A \wedge B$. An excellent introductory Logic text that uses sequent calculi is [Lem65].

Both Natural Deduction and Sequent Calculi were introduced by Gerhard Gentzen in the 1930's. I am unable to say much about *sequent calculi*, because I don't understand them. This will change when I have time to read more of Curry's *Foundations of Mathematical Logic* [Cur63]. This is, according to Robert K. Meyer, the best Logic book ever written. Curry is very difficult though. Perhaps [Res00, Chapt. 6] would be a better place to start. Sequent calculi are the hot topic in research and are important in computer science, and automated deduction.

Semantic Tableaux systems are a systematic search for a counter-model. That is, a valuation that satisfies the premises and the negated conclusion[3]. This is acheived by breaking down formulae into their component parts.

For example, let the *initial set* $I = \Gamma \cup \{\neg C\}$ of formulae contain $A \wedge B$. A valuation will satisfy $A \wedge B$ iff it satisfies A and B (assuming a reasonably sensible logic). We can therefore replace $A \wedge B$ by A and B. If I contains $C \vee D$, then there are two classes (not necessarily disjoint) of valuation that could satisfy the initial set: those that satisfy C; and those that satisfy D. We therefore create two sets of formulae, neither with $C \vee D$, one with C, the other with D. If a valuation satisfies either of these sets, it will satisfy I.

The procedure continues in this way until no more rules can be applied, because (in the the propositional case) each set of formulae contains only negated and unnegated propositional variables.

For example, we might be left with $\{A, \neg B\}$. Clearly the (classical) valuations that satisfy this set are those $\overline{v}$ where $\overline{v}(A) = True$ and $\overline{v}(B) = False$. The rules by which this set is obtained ensure that such a $\overline{v}$ will also satisfy the initial set. This valuation is a counter-model to the argument.

The sets of formulae might end up containing a propositional variable and its negation, eg $\{A, \neg A, \dots\}$. This set is clearly (classically) unsatisfiable. If all the resulting sets of formulae are unsatisfiable, then the initial set was too. This means that no countermodel exists for the argument, so it is valid. That is, its conclusion is a consequence of its premises. Smullyan [Smu68] introduced semantic tableaux. The *tree* proof method is an example of sematic tableaux; they are used in the introductory text [Jef91].

Given the huge range of deductive systems, our notion of deductive con-

[3]Since this is not always equivalent to satisfying the premises yet failing to satisfy the conclusion, there are logics for which semantic tableaux systems are not possible.

sequence will have to rest on the following:

Not a Definition 4.1.2. Let D be a system of deduction. A sentence $C \in \mathcal{L}$ is a *deductive consequence under* D of a set of sentences $\Gamma \subseteq \mathcal{L}$ iff D says it is.

Those interested in *paraconsistent* deductive systems are refered to [Won98]. Wong's approach there is quite compatible with ours: he studies the deductive systems in terms of consequence relations and closure operators.

Chapter 5

Consequence

In 3.7.6 we defined semantic consequence (over $\mathcal{L}$) and in 4.1.2 we almost defined deductive consequence. Here we will define the more general notion of a *consequence relation*. We will see that consequence relations give rise to *closure operators*, and that topological closure operators are examples of them.

5.1 Consequence Relations

Definition 5.1.1. A *consequence relation* $\vdash$ on a set X is a relation (in the usual sense, not that of 3.4.1) from subsets of X to members of X. ie $\vdash \subseteq \mathcal{P}(X) \times X$.

Notation 5.1.2. We write $\Gamma \vdash C$ iff $\langle \Gamma, C \rangle \in \vdash$. We will write $A \vdash C$ for $\{A\} \vdash C$.

Example 5.1.3. In a topological space $\langle X, \mathcal{T} \rangle$, the relation "$\Gamma$ has C as a limit-point", for $\Gamma \subseteq X$, $C \in X$, is a consequence relation. Recall that C is a limit-point of Γ iff every open set containing C contains points from Γ other than C.

Both semantic and deductive consequence relations are consequence relations in the sense just given.

5.2 Closure Operators

Definitions equivalent to this one appear in [B&S00, p. 21, defn. 5.1] and [Van71, p. 75]. It is interesting to note that although these operators are better known from topology and algebra, they originated with Tarski's work on logical consequence. See [B&S00, §5], [Tar30] and [Tar36].

Definition 5.2.1. If we are given a set X, a mapping $Cl : \mathcal{P}(X) \to \mathcal{P}(X)$ is called a *closure operator* on X if, for $\Gamma, \Sigma \subseteq X$, it satisfies:

CO1:	$\Gamma \subseteq Cl(\Gamma)$	(extensive)
CO2:	$Cl(Cl(\Gamma)) = Cl(\Gamma)$	(idempotent)
CO3:	$\Gamma \subseteq \Sigma \Rightarrow Cl(\Gamma) \subseteq Cl(\Sigma)$	(isotone or monotone)

Definition 5.2.2. The *consequence operator* Cn induced by the consequence relation $\vdash$ over X is the function $Cn : \mathcal{P}(X) \to \mathcal{P}(X)$ given by

$$Cn(\Gamma) = \{C : \Gamma \vdash C\}$$

for all $\Gamma \subseteq X$.

Example 5.2.3. Given a topological space $\langle X, \mathcal{T} \rangle$ the *closure* of a set $\Gamma \subseteq X$ is the union of Γ and its limit points. That is, the topological closure is the consequence operator induced by the limit point relation.

In this example, the consequence operator is a closure operator. The obvious question to ask at this stage is, which consequence relations give rise to consequence operators that are also closure operators?

Proposition 5.2.4. *The consequence operator Cn induced by the consequence relation $\vdash$ over X is a closure operator iff $\vdash$ satisfies the following:*

CR1: $A \vdash A \;\; \forall A \in X$
CR2: *for any* $\Gamma, \Sigma \subseteq X$, *if* $\Gamma \vdash S \;\; \forall S \in \Sigma$ *and* $\Sigma \vdash A$ *then* $\Gamma \vdash A$
CR3: *if* $\Gamma \subseteq \Sigma \subseteq X$ *and* $\Gamma \vdash A$ *then* $\Sigma \vdash A$

The correspondence between the CO's of definition 5.2.1 and the CR's of proposition 5.2.4 are clear enough that no proof is needed.

Definition 5.2.5. If a consequence relation satisfies CR1, CR2 and CR3 from proposition 5.2.4, we will call it a *closure relation.*

Now, these are properties that most sensible consequence relations will have. Firstly, if A is true, then A is true, so we are more interested in consequence relations where $A \vdash A$ holds for all A.

Secondly, if C is a consequence of sentences Σ that are in turn all consequences of Γ, then C is a consequence of Γ. For example, if $\Gamma \vdash S$ and $S \vdash C$ then any sensible consequence relation would also have $\Gamma \vdash C$.

Lastly you would hope that adding a sentence to a set would not make any sentence cease to be a consequence of the enlarged set. We should note however that nonmonotonic logics are studied, especially in artificial intelligence (see [Bre91], and further references in [Won98]). They are used to represent the reasoning of an agent who holds defeasable assumptions. Consider the following standard example.

Example 5.2.6.

$$\begin{aligned} \text{Tweety is a bird.} &\vdash \text{Tweety can fly.} \\ \text{Tweety is an emu.} &\vdash \text{Tweety can not fly.} \end{aligned}$$

This seems a fair enough way of thinking, but it violates CR3 from proposition 5.2.4. It's fair enough, though not infallible, because most birds can fly, especially birds called Tweety. But we ought to have

$$\text{Tweety is an emu.} \vdash \text{Tweety is a bird.}$$

since emus are birds. In this context, we don't consider it possible for "Tweety can fly." and "Tweety can not fly." to both be true[1], yet emus exist, so we should have

$$\text{Tweety is an emu.} \nvdash \text{Tweety can fly.}$$

So not *all* sensible logics are monotonic, however we will have no more to say about nonmonotonic logics here. The semantic consequence relations we have defined in 3.7.6 do satisfy these conditions.

Theorem 5.2.7. *Semantic consequence relations are closure relations.*

Proof. Let $A \in \mathcal{L}$, $\Gamma, \Sigma \subseteq \mathcal{L}$, and L be a logic over $\mathcal{L}$. Let $\models$ denote the consequence relation of L.

[1] Although, a glance at [Won98] suggests to me that nonmonotonic paraconsistent logics do exist.

The sentence A is clearly satisfied by every admissible valuation that satisfies A, so we have $A \models A$.

Assume that for each $S \in \Sigma$ we have $\Gamma \models S$, and also that $\Sigma \models A$. Assume that $\overline{va}$ is an admissible valuation satisfying Γ (if none exists then $\Gamma \models A$ follows immediately). For each $S \in \Sigma$ we have that $\overline{va}$ satisfies S, because $\Gamma \models S$. Therefore $\overline{va}$ satisfies Σ. But $\Sigma \models A$, so $\overline{va}$ satisfies A as well. Hence we have $\Gamma \models A$.

Now assume that $\Gamma \subseteq \Sigma$ and that $\Gamma \models A$. Assume that $\overline{va}$ is an admissible valuation satisfying Σ (if none exists then $\Sigma \models A$ follows immediately). Now $\overline{va}$ satisfies everything in Γ since everything in Γ is in Σ, and because $\Gamma \models A$, we have that $\overline{va}$ also satisfies A. So $\Sigma \models A$. □

In most deductive systems, a proof of C from Γ is automatically a proof of C from $\Gamma \cup \Sigma$, so the consequence relations that they induce are monotonic.

Our non-definition 4.1.2 does not permit us to establish anything useful about deductive consequence relations in general. However, given a specific deductive system, proving whether or it gives us a closure ought to be straight forward.

Terminology 5.2.8. If a deductive consequence relation is a closure relation, we will call the closure operator obtained from it a *deductive closure*. Closure operators obtained from semantic consequence relations will be called *semantic closures*.

Sets of sentences that include all of their consequences are are important enough to get a special name.

Definition 5.2.9. Let $Th \subseteq \mathfrak{L}$ and Cn be a deductive or semantic closure operator over $\mathfrak{L}$, then Th is a *theory* under Cn iff $Th = Cn(Th)$.

Another way of putting this is to say that theories are *closed* under their consequence operators. This technical definition, we assert, is an ideal toward which anything seriously calling itself a "theory" ought to strive. For example it seems unlikely that literary "theorists" could present their work in this way. Even physicists would have trouble I suspect (Hilbert attempted to formulate an axiomatic physics, but was not successful).

Next we'll look at a property that ought to be enjoyed by all deductive closures, but not all semantic closures. Again, see [B&S00, defn. 5.4]

Definition 5.2.10. A closure operator Cl on the set X is an *algebraic closure operator*, if for every $\Gamma \subseteq X$

$$\text{CO4:} \quad Cl(\Gamma) = \bigcup\{Cl(\Gamma') : \Gamma' \subseteq \Gamma, \text{ and } \Gamma' \text{ is finite}\}$$

This is related to the notion of *compactness*. Some notes on this can be found on page 68. The equivalent condition on consequence relations is

$$\text{CR4:} \quad \text{if } \Gamma \vdash C \text{ then } \Gamma' \vdash C \text{ for some finite } \Gamma' \subseteq \Gamma$$

Requirements 5.2.11. Deductive closures must be algebraic closure operations.

If $\Gamma \vdash_A C$ then the deductive apparatus A owes us a proof of it. Proofs are finite, so the set $\Gamma' \subseteq \Gamma$ of sentences used in the proof is finite. This same proof then shows that $\Gamma' \vdash_A C$.

The following example shows that there are semantic closures that are not algebraic closure operators.

Example 5.2.12. Let the truth functional interpretation of $\mathfrak{L}$ be that of Classical Propositional logic, and let the admissible valuations be those whose universe is denumerable (countably infinite), and whose functional interpretation is surjective when restricted to the nullary functions symbols. (That is, everything has a name)

Let Γ denote the set of sentences $\{R_{1,1}n_1, R_{1,1}n_2, R_{1,1}n_3, \ldots\}$. Then

$$\Gamma \models (\forall x_1)R_{1,1}x_1$$

so

$$(\forall x_1)R_{1,1}x_1 \in Cl(\Gamma)$$

but for any finite subset $\Gamma' \subset \Gamma$

$$\Gamma' \not\models (\forall x_1)R_{1,1}x_1$$

therefore

$$(\forall x_1)R_{1,1}x_1 \notin \bigcup\{Cl(\Gamma') : \Gamma' \subseteq \Gamma \text{ and } \Gamma' \text{ is finite }\}$$

If a semantic consequence relation has property CR4, then its logic is said to have *finitary semantic entailment* (see [Van71, p. 36]). Since one of the main goals of logic is to find deductive systems with the same consequence relation as a given logical interpretation, we are very interested in which logical interpretations have it. However, this is not a question we will explore further here.

Chapter 6

Simple Examples

6.1 An Inconsistent Theory of Equality in Modulo 2 Arithmetic

In this section, we will exhibit a very simple inconsistent, nontrivial theory. It is based on the theory $RM3^2$ of Meyer and Mortensen in [M&M84].

Before getting into the details, let's consider a scenario where a theory like this might make some sense. Imagine that you are required to determine whether or not two numbers m and n are equal, but all that you know is whether they are odd or even. That is, you have $[m]_2$ and $[n]_2$: their residue classes modulo 2.

If you are given $[0]_2$ and $[1]_2$, you may confidently assert that $m \neq n$, and deny that $m = n$. If you are given $[0]_2$ and $[0]_2$ though, you would answer "maybe" to both of the questions: "$m = n$?" and "$m \neq n$?"

Now for the details. Let L be a logical interpretation of $\mathcal{L}$ where

(i) the truth values TV are the set $\{True, truish, False\}$ with $True$ and $truish$ designated

(ii) negation $\neg^{\mathcal{M}} : TV \to TV$ is given by table 6.1

(iii) the universe is $\mathbb{Z}_2$.

(iv) zero has a name: $n_1^{\mathcal{F}} = [0]_2 \in \mathbb{Z}_2$

(v) the successor function is represented by $f_{1,1}$, that is $f_{1,1}^{\mathcal{F}} : \mathbb{Z}_2 \to \mathbb{Z}_2$, is the function where $f_{1,1}^{\mathcal{F}}([i]_2) = [i+1]_2$

(vi) equality $=^{\mathcal{R}}$: $\mathbb{Z}_2 \times \mathbb{Z}_2 \to TV$ is defined as follows: $=^{\mathcal{R}} ([m]_2, [n]_2) =$ *truish* iff $[m]_2 = [n]_2$, *False* otherwise

v	$\neg^{\mathcal{M}}(v)$
$True$	$False$
$truish$	$truish$
$False$	$True$

Table 6.1: The Negation Function of the Relevant Logic R

Recall that while there are many[1] interpretations and hence valuations of our language $\mathcal{L}$, we are delineating a class of *admissible* valuations, and from this we obtain a semantic consequence relation. In this case, we are being very restrictive, even specifying the universe of the interpretation.

Note that the function symbols n_1 and $f_{1,1}$ have been elevated in status from non-logical to logical symbols, as our specification of admissible valuations constrains their denotation.

We will consider the set of valid sentences under this consequence relation. This is the smallest theory under this consequence relation, the bottom of the lattice of closed sets. Let

$$Th = \{C \in \mathcal{L} : \models C\}$$

Proposition 6.1.1. *The theory Th is inconsistent.*

Proof. Consider the sentence $n_1 = n_1$. Now

$$\begin{aligned} \overline{va}(n_1 = n_1) &= \; =^{\mathcal{R}} (n_1^{\mathcal{F}}, n_1^{\mathcal{F}}) \\ &= truish \end{aligned}$$

so $n_1 = n_1 \in Th$. But

$$\begin{aligned} \overline{va}(\neg(n_1 = n_1)) &= \neg^{\mathcal{M}}(\overline{va}(n_1 = n_1)) \\ &= \neg^{\mathcal{M}}(truish) \\ &= truish \end{aligned}$$

so $\neg n_1 = n_1 \in Th$ as well. Since Th contains a sentence and the negation of that sentence, it is inconsistent. □

[1]Are they a proper class?

Proposition 6.1.2. *The theory Th is non-trivial.*

Proof. Consider the sentence $f_{1,1}(n_1) = n_1$. Now

$$\begin{aligned} \overline{va}(f_{1,1}(n_1) = n_1) &= =^{\mathcal{R}} (f_{1,1}^{\mathcal{F}}(n_1^{\mathcal{F}}), n_1^{\mathcal{F}}) \\ &= =^{\mathcal{R}} ([1]_2, [0]_2) \\ &= False \end{aligned}$$

since $1 \neq 0 \pmod 2$. Therefore $f_{1,1}(n_1) = n_1 \notin Th$, hence Th is non-trivial. □

6.2 Fuzzy Logics

Fuzzy logic is taken more seriously by engineers and marketing people than it is by logicians. However, it can be presented in our semantic framework.

It is interesting in that it is a *very* many-valued logic: there are uncountably many truth values. We will also see that this is not the only odd thing about fuzzy logic, as Dr Mark Colyvan has pointed out. This material is drawn mainly from [Pri01, Chapt. 11].

Let $TV = [0,1]$, $\neg^{\mathcal{M}} x = 1 - x$, $x \wedge^{\mathcal{M}} y = min(x,y)$, $x \vee^{\mathcal{M}} y = max(x,y)$ and

$$x \supset^{\mathcal{M}} y = \begin{cases} 1 \text{ if } x \leq y \\ 1 - (x - y) \text{ otherwise} \end{cases}$$

This conditional demands an explanation. Put simply, it makes the conditional represent *truth preservation.* We think of a sentence B being *truer* than a sentence A, under a given valuation $\overline{va}$ when $\overline{va}(A) \leq \overline{va}(B)$. Consider the conditional $A \supset B$. If B is truer than A then reasoning from A to B preserves truth. If on the other hand, the B is less true than the A, we are losing truth. The formula

$$1 - (\overline{va}(A) - \overline{va}(B))$$

represents the amount of truth we are losing[2].

We have not specified a set of designated truth values. For reasons explained in Section 7.4, we will want the set to be closed upwards. Intuitively, if b is truer than a, and a is (in some sense) true, then b ought to be true as

[2]I am resisting the urge to put quotes around everthing here.

well. So we need a lower bound d on D, and we must specify whether or not $d \in D$.

Proposition 6.2.1. *For any fuzzy logic (as described), either* $\models A \wedge \neg A$ *or* $\not\models A \vee \neg A$.

Proof. Let $a =_{df} \overline{va}(A)$. Then $\overline{va}(A \wedge \neg A) = min(a, 1-a)$. Therefore if $\not\models A \wedge \neg A$ then $\frac{1}{2} \notin D$. But then $\overline{va}(A \vee \neg A) = max(a, 1-a) = \frac{1}{2}$ when $a = \frac{1}{2}$. Hence $\not\models A \vee \neg A$. $\square$

Chapter 7

Non-standard Analysis

Non-standard Analyisis (Robinson [Rob61], [Rob74]) is a consistent theory, so although it has aroused some controversy, it is pretty tame by our standards.

Our model however is different to the "standard" non-standard model[1] The work usually done by a congruence relation is here done by our interpretation of the equality symbol. This is a technique made possible by our extention of the notion of a model. We will compare our model with the usual one, and also with the internalised equalities of Fearnley-Sander and Stokes [FS&S97] at the end of the chapter.

Much of the chapter is devoted to *free ultrafilters*. In particular, we work through the requirements for the relations $=, \leq$ and the logical connectives $\neg, \vee, \wedge$ and $\supset$, eventually deciding that what we need is a set of designated truth values that form a free ultrafilter on the truth values. This material might sit better in a chapter on ultraproduct models in general, but that chapter did not materialise.

First, we'll outline the basic idea of non-standard analysis: infinitesimals.

7.1 Differentiation Done Differently

Recall that in standard analysis, we find the derivative $\frac{dy}{dx}$ of a differentiable function, say $y = f(x) = x^2$ as follows

[1]Perhaps I ought to call it Non-standard Non-standard Analysis.

$$
\begin{aligned}
\frac{dy}{dx} &= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0}\frac{(x+h)^2-x^2}{h} \\
&= \lim_{h\to 0}\frac{x^2+2xh+h^2-x^2}{h} \\
&= \lim_{h\to 0}\frac{2xh+h^2}{h} \\
&= \lim_{h\to 0} 2x+h \\
&= 2x
\end{aligned}
$$

where the limit is defined by $\lim_{x\to a} f(x) = L$ iff for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(a) - L| < \varepsilon$ whenever $|x - a| < \delta$. Intuitively, this is seen as the ability to make $f(x)$ as close as you like to L by making x close enough to a.

Non-standard analysis allows us to put x *infinitely* close to a, yet still not equal to it. We find the derivative by letting h be not a real number but an infinitesimal. That is, h is such that $0 < h < r$, $\forall r \in \mathbb{R}^+$. We then find the derivative this way

$$
\begin{aligned}
\frac{dy}{dx} &= \operatorname{st}\frac{f(x+h)-f(x)}{h} \\
&= \operatorname{st}\frac{(x+h)^2-x^2}{h} \\
&= \operatorname{st}\frac{x^2+2xh+h^2-x^2}{h} \\
&= \operatorname{st}\frac{2xh+h^2}{h} \\
&= \operatorname{st} 2x+h \\
&= 2x
\end{aligned}
$$

Where st z denotes the *standard part* of z, that is the one and only[2] real number infinitely close to z.

[2] I owe the reader existence and uniqueness proofs. Our study does not get that far though.

To express statements of real analysis in $\mathfrak{L}$, all we need to do is choose from our alphabet, a symbol of the appropriate kind and arity for each of the functions and relations used in analysis. Rather than actually doing this, we will assume that it has already been done, and use the usual symbols to stand for the formal symbols.

For example, we write

$$\mathrm{st}\,\frac{(x+h)^2 - x^2}{h}$$

for

$$f_{1,1}(f_{2,4}(f_{2,2}(f_{2,5}(f_{2,1}(x_1, x_2), n_2), f_{2,5}(x_1, n_2)), x_2))$$

by letting st stand for $f_{1,1}$, x for x_1, h for x_2, 2 for n_2 and $+, -, \times, \div$ and exponentiation for $f_{2,1}, f_{2,2}, f_{2,3}, f_{2,4}$ and $f_{2,5}$ respectively.

What we aim to do in the remainder of this chapter is to create a model that satisfies all the sentences of analysis, and also the following statement asserting the existence of infinitesimals

$$(\exists x)\ (x \neq 0 \wedge \mathrm{st}(x) = 0)$$

7.2 Mathematics Done Differently

The usual way of describing the model that we need would be to say that it is a linearly ordered field that embeds $\langle \mathbb{R}, +, \times, \leq \rangle$. The usual conception of a mathematical structure does not explicitly include the truth values, or the logical operators. Rather, "classical" logic is silently assumed. So when one speaks of a *field*, one means a universe with a non-logical interpretation that satisfies the field axioms under the classical two valued logical interpretation.

We will call *any* model (see 3.7.5) that satisfies the field axioms, a field. And similarly for semigroups, rings and all the rest.

This is a rather bold move, that probably has implications that are either profound or unacceptable. Ring theory for example is divided into commutative and non-commutative parts. Maybe in the future, some ring theorists will focus on consistent rings, and others on inconsistent rings.

All that we can hope to accomplish here is to show that this move allows an interesting new approach to a known mathematical structure.

The model we will develop in the remainder of this chapter is a field in our sense, but not in the usual sense. We will see however that the usual non-standard model can be recovered from it in a fairly obvious way.

7.3 The Universe: Real Sequences

Let $\mathbb{R}^{\mathbb{N}}$ denote the set of sequences of real numbers $\langle x_1, x_2, \ldots \rangle$. We will sometimes write $\langle x_i \rangle$ for such a sequence. The universe U of our model will be $\mathbb{R}^{\mathbb{N}}$.

We define functions on $\mathbb{R}^{\mathbb{N}}$ pointwise in terms of their counterparts on $\mathbb{R}$. For example we define addition and multiplication pointwise on $\mathbb{R}^{\mathbb{N}}$ by setting

$$
\begin{array}{rcl}
\langle x_1, x_2, \ldots \rangle +^{\mathcal{F}} \langle y_1, y_2, \ldots \rangle & =_{df} & \langle x_1 + y_1, x_2 + y_2, \ldots \rangle \\
\langle x_1, x_2, \ldots \rangle \times^{\mathcal{F}} \langle y_1, y_2, \ldots \rangle & =_{df} & \langle x_1 \times y_1, x_2 \times y_2, \ldots \rangle
\end{array}
$$

7.4 Pointwise Truth

We have defined our operations on the real sequences pointwise, in terms of their counterparts on the real numbers. Relations (in the usual sense) do not lend themselves to this treatment, because given a pair of sequences, and a relation on the reals, you get not a True/False answer, but a sequence of them.

Relations in our (3.4.1) sense however, are more adaptable. Let $R: U^n \to \{True, False\}$ be an n-ary relation on some universe U for the truth values $\{True, False\}$. Then $R': (U^{\mathbb{N}})^n \to \{True, False\}^{\mathbb{N}}$ defined by

$$R'(\langle u_1^1, u_2^1, \ldots \rangle, \ldots, \langle u_1^n, u_2^n, \ldots \rangle) = \langle R(u_1^1, \ldots, u_1^n), R(u_2^1, \ldots, u_2^n), \ldots \rangle$$

is an n-ary relation on $U^{\mathbb{N}}$ for the truth values $\{True, False\}^{\mathbb{N}}$. Luckily, we are only interested in the binary operations $=$ and $\leq$, so we can tidy up our notation somewhat. The denotation of the equality symbol is given by

$$
\begin{array}{rcl}
\overline{va}(a = b) & = & \langle a_i \rangle =^{\mathcal{R}} \langle b_i \rangle \\
& =_{df} & \langle a_1 = b_1, a_2 = b_2, \ldots \rangle \\
& \in & \{True, False\}^{\mathbb{N}}
\end{array}
$$

where a and b are terms and $\overline{va}(a) = \langle a_i \rangle$, $\overline{va}(b) = \langle b_i \rangle$.

The relation $\leq^{\mathcal{R}}$ is defined similarly in terms of the usual $\leq$ on the real numbers.

The truth functions over these truth values can also be defined pointwise, in terms of their classical counterparts. For example conjunction is given by

$$\langle s_1, s_2, \dots \rangle \wedge^{\mathcal{M}} \langle u_1, u_2, \dots \rangle =_{df} \langle s_1 \wedge u_1, s_2 \wedge u_2, \dots \rangle$$

where $s_i, u_i \in \{True, False\}$ for all i, and the unsuperscripted $\wedge$ denotes classical two valued conjunction.

Note that this logical interpretation $\mathcal{M}$ is a Boolean algebra, as it is a power of the two valued Boolean algebra. Its order relation $\preceq$ is given[3] by

$$s = \langle s_i \rangle \preceq_{df} \langle w_i \rangle = w \text{ iff } s_k = True \Rightarrow w_k = True$$

It seems reasonable to think of w as "at least as true as" s when this holds, since w is $True$ in at least all the places where s is.

Will these interpretations of the logical connectives work anything like the way we expect? That is, will the consequence relation we obtain work the way we expect with respect to the logical connectives? Here are some[4] properties that we would like our connectives to have.

Requirements 7.4.1.

$$\varphi \wedge \psi \models \varphi \quad \text{and} \quad \varphi \wedge \psi \models \psi \tag{7.1}$$
$$\varphi, \psi \quad \models \quad \varphi \wedge \psi \tag{7.2}$$
$$\varphi \models \varphi \vee \psi \quad \text{and} \quad \psi \models \varphi \vee \psi \tag{7.3}$$
$$\varphi \vee \psi \models \varphi \quad \text{or} \quad \varphi \vee \psi \models \psi \tag{7.4}$$
$$\overline{va}(\neg\varphi) \in D \quad \text{iff} \quad \overline{va}(\varphi) \notin D \tag{7.5}$$
$$\varphi, \varphi \supset \psi \quad \models \quad \psi \tag{7.6}$$

for all sentences $\varphi, \psi \in \mathcal{L}$ and all admissible valuations $\overline{va}$.

That is we want the valuations to "respect" conjunction (7.1, 7.2) and disjunction (7.3, 7.4); to have "exclusion negation" (7.5, see [Van71, p. 37]); and admit detatchment (7.6) (ie Ackerman's γ rule, [F&M92]).

Note that logics which fail one or more of these requirements have been seriously proposed[5]

[3] Since I've already defined the meet and join of the lattice, I ought to prove that this is the order, not just assert it. Sorry.

[4] This list is not meant to be complete in any important sense. It turns out to be enough to motivate ultrafilters. There probably is some important sense concerning our need to satisfy the sentences of real analysis, in which it should be complete though.

[5] A catalogue of Logical principles, philosophical arguments for and against them, and the logics that satisfy and fail them would be very helpful in Logical work. Anyone with some funding who agrees should feel free to contact the author.

7.5 What Shall We Designate?

Our only means of controlling these things is by our choice of designated truth values. Can we choose $D \subseteq \{True, False\}^{\mathbb{N}}$ such that requirements 7.1 - 7.6 are satisfied?

All the results that follow in this section are specific to logical interpretations that are products of the two valued Boolean algebra. It may be possible to generalise some of them.

Definition 7.5.1. A *cone* of a set S partially ordered by $\preceq$, is a subset $C \subseteq S$ such that if $c \in C$ and $c \preceq d$ then $d \in C$. A *proper* cone is a cone $C \neq S$.

One might say that cones are "closed upwards". The next proposition shows that this is what we need in order to have the requirement 7.1.

Proposition 7.5.2. *The sentences φ and ψ are satisfied whenever $\varphi \wedge \psi$ is, iff D is a cone.*

Proof. Assume that D is not a cone. Then we have $p, s \in TV$ such that $p \preceq s$ and $p \in D$ but $s \notin D$. Let $\overline{va}(\varphi) = p$ and $\overline{va}(\psi) = s$. Then $\overline{va}(\varphi \wedge \psi) = p$, so the conjunction $\varphi \wedge \psi$ is satisfied by $\overline{va}$ but the conjunct ψ is not. That is $\varphi \wedge \psi \not\models \psi$.

If D is a cone, and $\overline{va}$ satisfies $\varphi \wedge \psi$ then because TV is a lattice we have

$$D \ni \overline{va}(\varphi \wedge \psi) = \overline{va}(\varphi) \wedge^{\mathcal{M}} \overline{va}(\psi) \preceq \overline{va}(\psi)$$

but as D is a cone, $\overline{va}(\psi)$ is in D too. Hence $\varphi \wedge \psi \models \psi$. Similarly for φ. $\square$

Now let's see what we need for requirement 7.2. We defined $\wedge^{\mathcal{M}}$ so that $\overline{va}(\varphi \wedge \psi) = \overline{va}(\varphi) \wedge^{\mathcal{M}} \overline{va}(\psi)$ and therefore we need D to be closed under $\wedge^{\mathcal{M}}$ in order for $\varphi \wedge \psi$ to be satisfied whenever φ and ψ are.

Definition 7.5.3. A *filter* on a lattice L is a cone F such that if $f, g \in F$ then $f \wedge g \in F$. A *proper* filter is a filter $F \neq L$.

We'll gather our progress so far into the following proposition:

Proposition 7.5.4. *Requirements 7.1 and 7.2 are satisfied iff D is a filter.*

So long as D is a cone, we can get some useful intuition by thinking of the lattice and reading $\models$ as $\preceq$. This makes it easier to sort out requirement 7.3.

Proposition 7.5.5. *A sentence $\varphi \vee \psi$ is satisfied whenever φ or ψ is, iff D is a cone.*

Proof. We know that $\overline{va}(\varphi) \preceq \overline{va}(\psi) \vee^{\mathcal{M}} \overline{va}(\psi)$, so wave your hands around, and we have $\varphi \models \varphi \vee \psi$. Similarly for ψ. It's dual to 7.5.2 $\square$

Notation 7.5.6. We will write b' for the complement of b in a Boolean algebra (complemented distributive lattice).

Definition 7.5.7. An *ultrafilter* on a Boolean algebra B is a filter $U \subseteq B$ where each $b \in B$ is such that $b \notin U$ iff $b' \in U$.

Proposition 7.5.8. *If TV is a Boolean algebra and $D \subseteq TV$ is an ultrafilter, then each $\overline{va}$ that satisfies $\varphi \vee \psi$ also satisfies either φ or ψ.*

Proof. Let $\overline{va}(\varphi \vee \psi) \in D$. This is equal to $\overline{va}(\varphi) \vee^{\mathcal{M}} \overline{va}(\psi) \in D$ because of the way we defined $\overline{va}$. But as TV is a Boolean algebra, we have

$$\overline{va}(\varphi) \vee^{\mathcal{M}} \overline{va}(\psi) = (\overline{va}(\varphi)' \wedge^{\mathcal{M}} \overline{va}(\psi)')'$$

by DeMorgan's Laws. Since D is an ultrafilter, $\overline{va}(\varphi)' \wedge^{\mathcal{M}} \overline{va}(\psi)' \notin D$. The set D is closed under $\wedge^{\mathcal{M}}$, therefore either $\overline{va}(\varphi)'$ or $\overline{va}(\psi)'$ is not in D. Assume (without loss of generality) that $\overline{va}(\varphi)' \notin D$. Then as D is an ultrafilter we have $\overline{va}(\varphi) \in D$. $\square$

None of the weaker properties we have defined will do. This seems a bit surprising, so we will have a more careful look. Consider these two counterexamples. In each case $a \vee b \in D$ but $a, b \notin D$. The designated values are indicated by the heavier dots. Firstly, in figure 7.1, we have a proper cone that is not a filter, which shows that D must be a filter for 7.4 to hold. Secondly, figure 7.2 is a filter on a Boolean algebra that is not an ultrafilter.

Now as our TV *is* a Boolean algebra, and because of the way we defined $\overline{va}$, we have that

$$\overline{va}(\varphi)' = \overline{va}(\neg\varphi)$$

This gives us requirement 7.5:

Proposition 7.5.9. $\overline{va}(\neg\varphi) \in D \Leftrightarrow \overline{va}(\varphi) \notin D$ *iff D is an ultrafilter.*

Now each truth projection $\langle v_1, v_2, \dots \rangle \mapsto v_k$ is a $\supset$, $\neg$ and $\vee$ homomorphism. Therefore, because $\varphi \supset \psi$ is equivalent to $\neg\varphi \vee \psi$ in classical logic, this equivalence also holds in our logical interpretation $\mathcal{M}$, regardless of which truth values we end up designating. This gives us what we need for requirement 7.6.

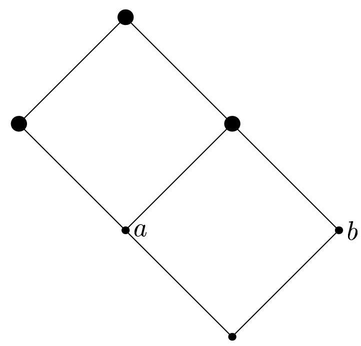

Figure 7.1: A proper cone that is not a filter, where $a \vee b$ is designated, but neither a nor b are.

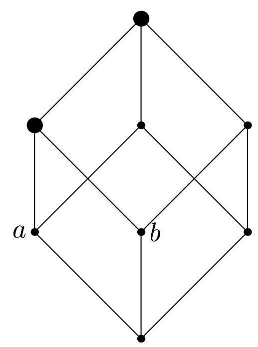

Figure 7.2: A filter on a Boolean algebra, that is not an ultrafilter, where $a \vee b$ is designated, but neither a nor b are. .

Proposition 7.5.10. *If TV is a Boolean algebra and D an ultrafilter, then*

$$\varphi, \varphi \supset \psi \models \psi$$

Proof. Let $\overline{va}$ satisfy $\varphi, \neg\varphi \vee \psi$. By 7.5.8 it satisfies $\neg\varphi$ or ψ. But it doesn't satisfy $\neg\varphi$, because of 7.5.9. Therefore $\overline{va}$ satisfies ψ. □

Now we have worked out what we need to make our logical connectives work properly[6]. Next we will investigate how to make our logical relations behave nicely. We won't assume that D satisfies any of the properties we

[6]In so far as "properly" is defined by the requirements 7.4.1.

have discussed in this section, rather we will invoke them only when they are needed.

7.6 Order and Equality: What We Need

Will our choice of designated values be constrained by our requirements for the relations $=$ and $\leq$?

In the following, we employ an innocent extention of our interpretive superscripts. We will write $a^{\mathcal{F}}$ for the denotation of a term a, even if a is not a simple name, but a compound term such as $f_{1,1}(n_1)$. It is innocent because the terms are literally a *term algebra*, and the interpretation function ι (defn. 3.3.1 page 18), that the superscript $\mathcal{F}$ is short for, is a homomorphism. Also, note that the "=" sign is used sometimes as a formal symbol of our language $\mathcal{L}$, and sometimes in the conventional way, to denote the relation "... is identical with ...". It's clear enough which is which if you take care.

It might seem natural to count equations as holding only when the terms on either side of the "=" denote the same object in the universe.

$$\models a = b \text{ iff } a^{\mathcal{F}} = b^{\mathcal{F}}$$

which is equivalent to

$$\overline{va}(a = b) \in D \text{ iff } a^{\mathcal{F}} = b^{\mathcal{F}}$$

but as the admissible valuation functions $\overline{va}$ are already determined, we can only satisfy this by choosing designated truth values. In this case, the only designated truth value would be

$$\langle True, True, True, \ldots \rangle$$

We will amuse ourselves by calling this logical interpretation the *naive interpretation*. Now, we must have $\models a = b$ whenever $a^{\mathcal{F}} = b^{\mathcal{F}}$, otherwise $\not\models a = a$ which is not acceptable as we wish to satisfy all the sentences of real analysis. But, as we shall see, we also need $\models a = b$ in some cases where $a^{\mathcal{F}} \neq b^{\mathcal{F}}$.

We need to take care though. If for example

$$\top_1^{\mathcal{M}} = \langle True, False, False, \ldots \rangle \in D$$

but

$$\top_2^{\mathcal{M}} = \langle True, True, False, False, \dots \rangle \notin D$$

then we would have

$$\langle 1, 1, \dots \rangle =^{\mathcal{R}} \langle 1, 2, 3, \dots \rangle = \top_1^{\mathcal{M}} \in D$$

and

$$\langle 1, 1, \dots \rangle =^{\mathcal{R}} \langle 1, 0, 0, \dots \rangle = \top_1^{\mathcal{M}} \in D$$

but

$$\langle 1, 1, \dots \rangle +^{\mathcal{F}} \langle 1, 1, \dots \rangle =^{\mathcal{R}} \langle 1, 2, 3, \dots \rangle +^{\mathcal{F}} \langle 1, 0, 0, \dots \rangle = \top_2^{\mathcal{M}} \notin D$$

Given the obvious denotations for the formal names a, b, c and d then, we would have

$$\models a = b \text{ and } \models c = d$$

but

$$\not\models a + c = b + d$$

That is, $=^{\mathcal{R}}$ is not compatible with $+^{\mathcal{R}}$.

Note that in this example, D is not a cone (defn. 7.5.3).

The naive interpepation gives us a commutative ring with zero $\langle 0, 0, \dots \rangle$ (for which we will write $\langle 0 \rangle$) and one $\langle 1, 1, \dots \rangle = \langle 1 \rangle$.

But under the naive interpretation we also have zero divisors, since for example

$$\begin{aligned} \langle 1, 0, 1, 0, \dots \rangle &\neq \langle 0 \rangle \\ \langle 0, 1, 0, 1, \dots \rangle &\neq \langle 0 \rangle, \text{ but} \\ \langle 1, 0, 1, 0, \dots \rangle \times \langle 0, 1, 0, 1, \dots \rangle &= \langle 0 \rangle \end{aligned}$$

To eliminate zero divisors from the theory, we require that one of each pair of zero-divisors is equal to zero. That is, whenever $\models ab = 0$ we must have $\models a = 0$ or $\models b = 0$.

That still would not give us a field, because sequences that contain 0 lack a multiplicative inverse. If $\langle a_i \rangle \times \langle b_i \rangle = \langle 1 \rangle$ and $a_k = 0$ then $0 \times b_k = 1$ which can't happen with real numbers.

Note that any sequence $\langle a_i \rangle$ where $a_i \neq 0 \ \forall i \in \mathbb{N}$, does have a multiplicative inverse $\langle a_i^{-1} \rangle$. Call these sequences, *nowhere-zero*. If the equality relation $=^{\mathcal{R}}$ relates each 0-containing sequence to either $\langle 0 \rangle$ or a nowhere-zero sequence, then the structure will have[7] multiplicative inverses for all non-zero elements.

To see this, let $\langle a_i \rangle$ contain 0, $\langle b_i \rangle$ be nowhere-zero, and let the formal names a and b denote $\langle a_i \rangle$ and $\langle b_i \rangle$ respectively.

We have

$$\models b \times b^{-1} = 1$$

so if

$$\models a = b$$

and the theory is closed under the substitution of equals[8], then

$$\models a \times b^{-1} = 1$$

that is, we have our multiplicative inverse for a.

We need to choose designated values such that $\leq$ is a linear order. The naive interpretation will not give us a linear order though. Let a denote the sequence $\langle 1, 0, 0, \ldots \rangle$ and b the sequence $\langle 0, 1, 0, 0, \ldots \rangle$ then because

$$\begin{array}{rcl} \langle 1, 0, 0, \ldots \rangle \leq^{\mathcal{R}} \langle 0, 1, 0, 0, \ldots \rangle & = & \langle False, True, True, \ldots \rangle \\ \langle 0, 1, 0, 0, \ldots \rangle \leq^{\mathcal{R}} \langle 1, 0, 0, \ldots \rangle & = & \langle True, False, True, True, \ldots \rangle \end{array}$$

we have

$$\begin{array}{rcl} \not\models a & \leq & b \\ \not\models b & \leq & a \end{array}$$

We require at least one of these to be satisfied. In other words, we need at least one of $\langle False, True, True, \ldots \rangle$ and $\langle True, False, True, True, \ldots \rangle$ to be designated. In fact, we shall see that they both are, and as a result we will also need $\models a = b$, that is $\langle False, False, True, True, \ldots \rangle$ must be designated.

To summarise then:

[7] This is a case where our way of doing maths is different. Although $\models a = b$ this does *not* mean that a and b denote the same object in the universe. The relevant sentences about existence of inverses are satisfied however.

[8] I think this is equivalent to equality being a congruence, which we prove at 7.7.2.

Requirements 7.6.1. We need to choose a set of designated truth values $D \subseteq TV$, such that the following are satisfied:

(i) equality is a congruence

(ii) at least one of each pair of zero-divisors must equal zero (in order to eliminate zero divisors)

(iii) each zero-containing sequence must equal either a nowhere-zero sequence or zero (in order to get $\times$-inverses)

(iv) $\leq$ is a linear order

(v) the theorems of real analysis take designated values (if this isn't a consequence of the above)

(vi) we have valid sentences asserting the existence of infinitesimals

7.7 Filters and Congruence

We have already determined that $\langle True\rangle$ must be designated in order for $\models a = a$ to hold for all names a, ie for equality to be reflexive.

Since

$$\begin{aligned}\langle a_i\rangle =^{\mathcal{R}} \langle b_i\rangle &= \langle a_i = b_i\rangle \\ &= \langle b_i = a_i\rangle \\ &= \langle b_i\rangle =^{\mathcal{R}} \langle a_i\rangle\end{aligned}$$

we don't have to do anything special to make equality symmetric.

We also require transitivity,

$$\text{if } \models a = b \text{ and } \models b = c \text{ then } \models a = c$$

That is

$$\text{if } \langle a_i\rangle =^{\mathcal{R}} \langle b_i\rangle \in D \text{ and } \langle b_i\rangle =^{\mathcal{R}} \langle c_i\rangle \in D \text{ then } \langle a_i\rangle =^{\mathcal{R}} \langle c_i\rangle \in D$$

Now if $\langle a_i = b_i\rangle \in D$ and $\langle b_i = c_i\rangle \in D$, we would like

$$\langle a_i = b_i\rangle \wedge^{\mathcal{M}} \langle b_i = c_i\rangle = \langle a_i = b_i \wedge b_i = c_i\rangle$$

to be in D too. That is, we want requirement 7.2 from page 45 to hold. But $a_i = b_i \wedge b_i = c_i \Rightarrow a_i = c_i$ and therefore

$$\langle a_i = b_i \wedge b_i = c_i \rangle \preceq \langle a_i = c_i \rangle$$

We have therefore shown that

Proposition 7.7.1. *If D is a filter, then $=$ is an equivalence relation.*

Paul Hunter, my bright young classmate from 2000, took an interest in these matters, and wrote [Hun00], where he proved the converse: if $=$ is an equivalence then D is a filter.

Proposition 7.7.2. *If D is a filter, then $=$ is a $+$ and $\times$-congruence.*

Proof. Let $\models a = b$, ie $\langle a_i = b_i \rangle \in D$. Now $\langle a_i \rangle \times^{\mathcal{F}} \langle c_i \rangle = \langle a_i c_i \rangle$ and $\langle b_i \rangle \times^{\mathcal{F}} \langle c_i \rangle = \langle b_i c_i \rangle$. But $a_i = b_i \Rightarrow a_i c_i = b_i c_i$, so

$$D \ni \langle a_i = b_i \rangle \preceq \langle a_i c_i = b_i c_i \rangle$$

and since D is closed upwards, we have this last truth value designated too, ie $\models ac = bc$. The argument for $+$ is similar. □

It is possible to show that the equivalence induced by a filter is a congruence for *every* pointwise defined function (and hence our relations). Again, see [Hun00]. This makes the presentation more complicated however, and it is a consequence of the *transfer principle* which, tragically, is not to be found in this study. See [H&L85, I.5, I.15].

7.8 Ultrafilters

In this section, we will see that a congruence induced by an ultrafilter (defn 7.5.7) will satisfy many of our requirements.

Zero-Divisors

The following proposition shows that if the designated values form an ultrafilter, the theory will be free from zero-divisors.

Proposition 7.8.1. *Let D be an ultrafilter on TV. If $\models ab = 0$ for some terms a and b, then $\models a = 0$ or $\models b = 0$.*

Proof. Let $\models ab = 0$. Then $\langle a_i b_i = 0\rangle \in D$. Assume $\not\models a = 0$, ie $\langle a_i = 0\rangle \notin D$. Since D is an ultrafilter, $\langle a_i = 0\rangle' = \langle a_i \neq 0\rangle \in D$. But if $a_k b_k = 0$ and $a_k \neq 0$ then $b_k = 0$, so we have

$$\langle a_i b_i = 0\rangle \wedge^{\mathcal{M}} \langle a_i \neq 0\rangle = \langle a_i b_i = 0 \wedge a_i \neq 0\rangle \preceq \langle b_i = 0\rangle$$

The left-hand-side of the $\preceq$ is in D because it is the meet of two values in D. Hence the right-hand-side is also in D, because D is closed upward. That is we have $\models b = 0$ as required. $\square$

Multiplicative Inverses

As we mentioned earlier, to ensure that everything has a multiplicative inverse, we must show that each zero-containing sequence is equal to either a nowhere-zero sequence or to the zero sequence[9].

Proposition 7.8.2. *Let D be an ultrafilter. If $\langle a_i\rangle$ is such that $a_k = 0$ for some k, then either*

(i) $\models a = 0$, *or*

(ii) there is a sequence $\langle b_i\rangle$ where $b_i \neq 0$ for all i, and if $b^{\mathcal{F}} = \langle b_i\rangle$ then[10] $\models a = b$

Proof. Let the function $u : \mathbb{N} \to \mathbb{N}$ be given by

$$u(x) = \begin{cases} 1 & \text{if } x = 0 \\ x & \text{otherwise} \end{cases}$$

then $u(a_i) = a_i$ unless $a_i = 0$. That is

$$\langle a_i = u(a_i)\rangle = \langle a_i \neq 0\rangle = \langle a_i = 0\rangle'$$

Now, if $\not\models a = 0$, ie $\langle a_i = 0\rangle \notin D$, then, since D is an ultrafilter, $\langle a_i = u(a_i)\rangle \in D$, hence $\langle u(a_i)\rangle$ is the nowhere-zero $\langle b_i\rangle$ that we require. $\square$

[9] Not both, otherwise everthing equals zero.

[10] We need to take care not to overstate our case here. Most of the objects in the universe have no name. We also have $\models (\exists y)a = y$ where y "represents" a sequence with the properties we ascribe to $\langle b_i\rangle$. But this kind of handwaving can only go in a footnote.

Linear Ordering

Proposition 7.8.3. *If D is an ultrafilter then $\leq$ is a linear ordering.*

Proof. Assume $\not\models a \leq b$. That is, $\langle a_i \leq b_i \rangle \notin D$. As D is an ultrafilter, we have

$$D \ni \langle a_i \leq b_i \rangle' = \langle b_i < a_i \rangle \preceq \langle b_i \leq a_i \rangle$$

and since D is closed upward, this last truth value is also designated. That is $\models b \leq a$. $\square$

7.9 Free Ultrafilters and Infinitesimals

All that remains is to show that we have infinitesimals. Oh yes, and that we model the reals.

We have been working for many pages now, toward a rigourous definition of "true enough", that will satisfy requirements 7.4.1 and 7.6.1. We are almost there, and it is time for the readers' patience to be rewarded with an intuitive version. True enough is *finitely false.* That is, we will designate all the truth values that are false in only finitely many places.

Definition 7.9.1. An element $\langle u_i \rangle \in TV$ is said to be *finitely false* iff

$$|\{i : u_i = False\}| \in \mathbb{N}$$

Similarly for *finitely true.*

Definition 7.9.2. An ultrafilter on TV is a *free ultrafilter* if it contains all finitely false elements.

It's not obvious that free ultrafilters exist. It is shown in [H&L85, Appendix] that their existence follows from Zorn's lemma (ie from the axiom of choice).

Hunter [Hun00] has shown that if D is an ultrafilter that is not free, then our model is isomorphic to the reals. My adaptation of Paul's proofs follow.

Note that the atoms of lattice TV are the elements that are true in exactly one place. That is they have the form

$$\langle False, False, \ldots, True, False, \ldots \rangle$$

Also, given any subset S of TV, we can generate a cone by including everything β such that $s \preceq \beta$ for some $s \in S$.

Definition 7.9.3. We will call the cone generated by a set containing a single atom, an *atomic cone.*

It's not too hard to see that an atomic cone consists of all the truth values with $True$ in the coordinate where the generating atom has its one and only $True$. The meet of two of these things will also have $True$ at that coordinate. Everything that is out of the atomic cone has $False$ there, so it's complement is in. That is

Proposition 7.9.4. *Each atomic cone in TV is an ultrafilter.*

Atomic cones are not *free* ultrafilters though, since the complement of the generating atom

$$\langle False, False, \ldots, True, False, \ldots \rangle' = \langle True, True, \ldots, False, True, \ldots \rangle$$

is finitely false, but is not in the atomic cone. In fact,

Proposition 7.9.5. *Every non-free ultrafilter on TV is an atomic cone.*

Proof. Let D be an ultrafilter on TV, such that $\alpha \in TV$ is finitely true but $\alpha' \notin D$, ie α' is finitely false, and D is non-free. Now there are atoms $\alpha_1, \ldots, \alpha_k$ such that $\alpha = \alpha_1 \vee^{\mathcal{M}} \ldots \vee^{\mathcal{M}} \alpha_k$, and as TV is a Boolean algebra $\alpha' = \alpha_1' \wedge^{\mathcal{M}} \ldots \wedge^{\mathcal{M}} \alpha_k'$. If none of the α_i were in D, then, as D is an ultrafilter, all of the α_i' would be in D. Since D is closed under finite meet, α' would be in D too. But, by our assumption, it isn't. Therefore we have an atom $\alpha_j \in D$.

If there was some $\beta \in D$ where $\alpha_j \not\preceq \beta$, then β must be false in the one and only place where α_j is true. Therefore we would have $\alpha_j \wedge^{\mathcal{M}} \beta = \langle False \rangle \in D$ and hence $D = TV$, and that's no ultrafilter. So everything in D is $\preceq \alpha_j$, and as ultrafilters are closed upward, everything $\preceq \alpha_j$ is in D, ie $\beta \in D$ iff $\alpha_j \preceq \beta$. $\square$

So if D is a non-free ultrafilter, there will be some $d \in \mathbb{N}$ such that truth values will be designated iff they are true at d.

The sequences that will represent infinitesimals are those that converge to zero. The problem with non-free ultrafilters is that they make all of these sequences equal to zero.

Definition 7.9.6. A sequence $\langle a_i \rangle$ is *infinitesimal* iff $\langle |a_i| \leq r \rangle \in D$ for all $r \in \mathbb{R}^+$.

Proposition 7.9.7. *If D is a non-free ultrafilter, then zero is the only infinitesimal.*

Proof. Let D be a non-free ultrafilter, $\langle a_i \rangle$ an infinitesimal. Then $\langle |a_i| \leq r \rangle \in D$ for all $r \in \mathbb{R}^+$. Let $d \in \mathbb{N}$ be the true coordinate of the atom generating D. Then $|a_d| \leq r$ for all $r \in \mathbb{R}^+$, therefore $a_d = 0$, so $\langle a_i = 0 \rangle \in D$ ie $\models a = 0$. □

The standard part function $\mathrm{st}^{\mathcal{F}}$ is such that for finite[11] $\langle a_i \rangle$

$$\langle a_i \rangle - \mathrm{st}^{\mathcal{F}}(\langle a_i \rangle)$$

is infinitesimal. It follows that

$$\models (\forall x)\ \mathrm{st}(x) = x$$

ie, everthing is real.

We want nonzero infinitesimals. Consider the following sequence

$$\left\langle 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots \right\rangle$$

for which we will write $\langle \frac{1}{i} \rangle$. It is non-zero if anything is, because $\left\langle \frac{1}{i} = 0 \right\rangle = \langle False \rangle$. If we take $r \in \mathbb{R}^+$ then

$$\left\langle \frac{1}{i} \leq r \right\rangle = \langle False, False, \dots, True, True, \dots \rangle$$

which is finitely false. We therefore have

Proposition 7.9.8. *If D is a free ultrafilter, then our model contains infinitesimals*

This is not quite satisfactory though. We would like the theory itself to assert the existence of infinitesimals. We can't quantify over the reals there, only over the whole extended field.

We have already established that $\langle \frac{1}{i} \rangle$ is nonzero, and

$$\left\langle \frac{1}{i} \right\rangle -^{\mathcal{F}} \langle 0 \rangle = \left\langle \frac{1}{i} - 0 \right\rangle$$

[11] The reciprocals of infinitesimals are infinite, and have no standard part.

is infinitesimal, therefore

$$\mathrm{st}^{\mathcal{F}} \left\langle \frac{1}{i} \right\rangle = \langle 0 \rangle$$

We have finally reached the goal we set ourselves way back on page 43

Proposition 7.9.9. *If D is a free ultrafilter, then*

$$\models (\exists x)\ (x \neq 0 \wedge \mathrm{st}(x) = 0)$$

7.10 Comparisons

Our pointwise equality function $=^{\mathcal{R}}$ is an example of an *internalised equality* [FS&S97]. This constitutes evidence, depending on your point of view, either that internalised equalities are many-valued logics by stealth, or that many-valued logics are internalised equalities by stealth.

Let the relation (in the usual sense) σ on $\mathbb{R}^{\mathbb{N}}$ be given by

$$\langle a_i \rangle\ \sigma\ \langle b_i \rangle \text{ iff } \models a = b$$

Then the usual model of non-standard analysis is $\mathbb{R}^{\mathbb{N}}/\sigma$. The details are left as an exercise for the reader.

Appendix A

Towards an Inconsistent Solution to an Undecidable Problem

There is a fairly straight-forward problem in commutative group theory that is not decidable in Zermelo-Fraenkel (ZF) set theory. It is known as the *Whitehead problem*.

That is, neither the *Whitehead sentence*, nor its negation is a theorem of ZF. The problem is proved to be undecidable by giving two consistent extensions of ZF, one that satisfies the *Whitehead sentence*, and one that satisfies its negation. The commutative group theorists have responded to this situation by prefixing their theorems with the variant of ZF that the theorem is shown to hold in.

This practice, we claim, amounts to "mathematical pluralism". (Compare "logical pluralism", [B&R00]) They are saying that their theorems are kind-of-true, yet kind-of-false. Wouldn't it be good to embrace all this with an inconsistent yet nontrivial theory of abelian groups, in which the Whitehead sentence *and* its negation are satisfied? I haven't yet managed to do this, but this appendix outlines how it might work.

The idea for this application of inconsistent mathematics is due to my supervisor, Dr. Barry Gardner. The main reference is [Ekl76], from which some of the definitions and theorems are taken almost directly.

A.1 The Whitehead Problem

Definition A.1.1. A *free abelian group* is an abelian (commutative) group A generated by a set $X \subseteq A$ such that any map $\alpha : X \to B$ into an abelian group B, can be uniquely extended to a homomorphism $\overline{\alpha} : A \to B$.

Terminology A.1.2. We say that X *freely generates* A.

Free algebras are important because they can act as "representatives" of a type of algebra when proving that all the algebras of that type have a certain property.

Definition A.1.3. A surjective homomorphism of abelian groups $\pi : B \to A$ is said to *split* if there is a homomorphism $\rho : A \to B$ such that $\pi(\rho(a)) = a\ \forall a \in A$.

Terminology A.1.4. We call ρ a *splitting homomorphism for* π.

Theorem A.1.5. *An abelian group A is free iff every homomorphism onto A splits.*

Proof. Suppose that A is free and that $\pi : B \to A$ is surjective. If $\{X_i : i \in I\}$ freely generates A, choose $b_i \in B$ for each $i \in I$ such that $\pi(b_i) = x_i$. Since X freely generates A, there is exactly one homomorphism $\rho : A \to B$ such that $\rho(x_i) = b_i$ for each $i \in I$. Clearly ρ is a splitting homomorphism for π.

To prove the converse, consider an abelian group F freely generated by $X = \{x_a : a \in A\}$. Let $\pi : F \to A$ be the unique homomorphism such that $\pi(x_a) = a$ for all $a \in A$. By hypothesis, there is a splitting homomorphism $\rho : A \to F$ for π. Since ρ is injective, A is isomorphic to a subgroup of F; therefore, since subgroups of free abelian groups are free, we have that A is free. ☐

The Whitehead condition weakens "every homomorphism" to a special class of homomorphisms. Recall that the kernel $ker(\pi)$ of a group homomorphism $\pi : B \to A$ is the set of elements that π maps to the identity: $\{b : \pi(b) = e_A\}$.

Definition A.1.6. An abelian group A is called a *Whitehead group* (or *W-group*) iff each surjective homomorphisms $\pi : A \to B$ with $ker(\pi) \cong \mathbb{Z}$, splits.

Whitehead's problem is the following: *Is this weaker condition enough to ensure that an abelian group is free?* Or equivalently: *Are there any non-free Whitehead groups?*

A.2 Incompleteness and Undecidability

The Whitehead problem was proved to be undecidable in Zermelo-Fraenkel set theory by Shelah [She74] in 1974.

Zermelo-Fraenkel set theory is a formal theory, like those we have been studying in this work. The language is similar enough to our $\mathcal{L}$ that we can safely neglect the distinction. Its consequence relation, and hence its' closure operator, are induced by a Hilbert-style proof system. The theory is the closure of a set of axioms that are widely held to be *the* axioms of Mathematics.

Terminology A.2.1. We will denote the axioms of Zermelo-Fraenkel set theory, ZF, the consequence relation $\vdash$ and its closure operator Cn.

Now the famous incompleteness theorems of Gödel [Göd31] tell us that, assuming that $Cn(ZF)$ is consistent, there will be sentences $S \in \mathcal{L}$ such that $ZF \nvdash S$ and $ZF \nvdash \neg S$.

One such sentence is the "continuum hypothesis" (CH), which we can write informally as $\aleph_1 = 2^{\aleph_0}$. That is, ZF can give no answer to the following question: *Is there an infinite set larger than the set of integers, but smaller than the set of real numbers.* The continuum hypothesis says no, there is not.

Many mathematicians feel that there really are answers to questions like this[1], and that ZF is therefore inadequate as an axiomatisation of (all) mathematics. There is considerable debate about what axioms (if any) ought to be added to ZF to yield a "more" complete theory. Some of the candidates are:

CH the continuum hypothesis

V=L Gödel's axiom of constructibility

MA Martin's axiom

Each of these has been proved relatively consistent with ZF: you can take any one of these sentences, or their negations, add it to ZF, and the closure of that set under the usual consequence relation is consistent, so long as $Cn(ZF)$ is.

[1] The author isn't even sure whether there is an answer to the question "Are there answers to questions like the continuum hypothesis?"

Now, Gödels' axiom is stronger than the continuum hypothesis, that is

$$V = L, ZF \vdash CH$$

but Martin's axiom is *relatively independent*[2] of the continuum hypothesis: ie $MA, ZF \nvdash CH$ and $MA, ZF \nvdash \neg CH$.

There is one more sentence for which we need a name:

W "there is a non-free Whitehead group"

Theorem A.2.2. *$V = L, ZF \vdash W$ but $MA, \neg CH, ZF \vdash \neg W$*

Since this result was discovered, abelian group theorists have begun prefixing many of their theorems with the additional axioms they require. For example [Wes96]. Maybe there is no "one true Mathematics". In the remaining sections of this appendix, we take the position of someone who wants to accept *both* of these theories as (in some sense) true yet *denies* that (for example) "1=2".

A.3 An Inconsistent Non-Theory

Given a consistent theory Th, we can construct a map v_{Th} from the language $\mathcal{L}$ into the truth values $\{T, F, U\}$ (true, false, undecided) as follows:

$$v_{Th}(S) = \begin{cases} T \text{ if } S \in Th, \\ F \text{ if } \neg S \in Th, \\ U \text{ otherwise.} \end{cases}$$

These are the truth values of the "strong" Kleene Logic K_3 (see [Pri01, §7.3], [Kle52, §64], [Kle38]). The only designated value in this matrix is T. The propositional logic LP (see [Pri01, §7.4]) is the same, except that both T and U are designated. What if we just throw in all the undecidable sentences? Would that be the theory we are looking for?

Not really. This is just a set of sentences. It's not closed under a consequence relation, because we don't have a consequence relation. We have only a single valuation. This would not give a very interesting consequence relation (exercise: characterise the consequence relation given by a single admissible valuation) What we need is a (large) class of admissible valuations.

[2]Although this is a technical term, an intuitive idea will suffice for our purposes.

I haven't been able to complete this work yet, but the following thoughts seem to be of some value as a starting point.

In Chapter 7 our set of truth-values was a product of the usual set. We could do something similar here, using 3 copies of $\{True, False\}$, one for ZF and one for each of the extended theories. This would give us a lattice of 27 values. But many of these values could never occur, since whatever is $True$ in ZF is true in both of the extensions, and similarly for $False$. Only the sentences that are undecided by ZF can take different values in the extended theories. Eliminating these "useless" values, we obtain a sublattice of 11 values, with top element TTT, bottom FFF, and nine values in the middle.

This still does not give us a consequence relation though, just a more elaborate classification of sentences.

I think what is required is to move from *pointwise truth* to *pointwise consequence*. That is we designate subsets of the set of theories. There would be no point including ZF in this set, as every consequence under ZF is a consequence under either of the extensions. For example if we designate all the non-empty subsets, then $\Gamma \vdash_W C$ iff $\Gamma, V = L, ZF \vdash C$, or $\Gamma, \neg CH, MA, ZF \vdash C$.

The difficulty then is requirement 7.2 from page 45. We would have $\vdash_W W$ and $\vdash_W \neg W$, but $\not\vdash_W W \wedge \neg W$.

I think what is needed is to clarify the idea "product of models". Is the model presented in Chapter 7 a product of models? What else could it be? How does this idea relate to "pointwise consequence"?

It is a little distressing to have to leave this work in such a state of disarray. The author would welcome any communication on the subject.

Let me conclude with some of the things which seem desirable in any potential solution.

Requirements A.3.1. An inconsistent solution to the Whitehead problem should satisfy the following:

(i) We want a consequence relation $\vdash_W$ on $\mathfrak{L}$ such that

$$V = L, \neg CH, MA, ZF \vdash_W W \wedge \neg W$$

but the closure operator Cn_W induced by $\vdash_W$ does not trivialise these (inconsistent) premises.

That is

$$W \wedge \neg W \in Cn_W(V = L, \neg CH, MA, ZF) \neq \mathfrak{L}$$

(ii) We want the new theory to contain both of the consistent extensions to the Zermelo-Fraenkel theory.

$$Cn(V = L, ZF) \cup Cn(MA, \neg CH, ZF) \subseteq Cn_W(V = L, MA, \neg CH, ZF)$$

(iii) We would like the paraconsistent consequence relation Cn_W to be as "close as possible" to the usual one. Here are some properties that might suffice

$$\begin{aligned} Cn_W(V = L, ZF) &= Cn(V = L, ZF) \\ Cn_W(MA, \neg CH, ZF) &= Cn(MA, \neg CH, ZF) \\ Cn_W(V = L, MA, \neg CH, ZF) &= Cn_W(Cn(V = L, ZF), Cn(MA, \neg CH, ZF)) \end{aligned}$$

Appendix B

What Else?

Although this project has been underway for two years, I have not been able to accomplish very much of what I had hoped to do. Rather than completely abandon the outstanding items, I have gathered thoughts and references here.

It probably strikes the reader that I have developed a rather elaborate framework in this study, yet have not done very much with it. I hope that the thoughts collected here seem interesting, even exciting enough to warrant the little that I have done.

I also hope that this work will be continued. If I continue it, these notes will be helpful to refresh my memory. If others find my approach to this topic attractive, they may also benefit from my suggestions for its further development.

Here then is what amounts to a "brain dump". The reader is asked to forgive its shortcomings in style and organisation.

1. Introduction

   1.1 Consequences and possibility are closely linked. One way of seeing it, is that a sentence C is a consequence of a collection of sentences Γ if it is impossible for all the sentences in Γ to be true without C also being true. That is, consequence is determined by possibility.

   What then if it is impossible for all the sentences in Γ to be true together? Then every sentence is a consequence of Γ.

   What if we see the connection the other way round? We find the consequences of Γ by "extracting information" (whatever that means), then we decree that a collection of sentences that includes all its consequences (a theory) describes a *possibility*.

1.2 Logicians are not concerned with petty details like the speed of light, the existence of gods, or that diamonds are stronger than mushrooms. These are mere accidents of reality. What matters to us are the things that could not be otherwise. However, it seems that whenever anybody suggests that something could not be otherwise, somebody else denies it. Maybe everything is possible, and no consequences can be established without some background assumptions. To illustrate this, consider the consequences of the following sentenence:

   The laws of logic do not apply.

   (adapted from a similar example in [Pri01])

2. Language

   2.1 To keep everything satisfactory to Herr Hilbert, our alphabet ought to be finite. At least mine is countable. Enderton [End72, §2.8, p. 164] uses an uncountable alphabet.

   2.2 The chapter on language really ought to conclude with some examples of the formulae of the language, and their derivations. It is probably rather difficult to understand as it is.

3. Semantics

   3.1 On the matter of extensional versus intensional meaning: Do possibility constraints give us intentional meaning somehow? Or is this just a more refined form of extensional meaning? The phrases "has a heart", and "has kidneys", mean something different because it is *possible* for a creature to satisfy one and not the other. Does Ed Zalta's book [Zal88], address this?

   3.2 The relationship between a set of sentences, the models that satisfy them, and the sentences satisfied by all these models is an example of a *Galios correspondence*. See [Bir40, Chapter V Section 8]. Other examples include Zariski topologies [Har77], and, of course, Galois groups.

   3.3 I would like to have explored the connections with Universal Algebra much more. A propositional language, such as $\mathcal{S}$ is a free algebra generated by its propositional variables. How (and if) this extends to the full first order language $\mathcal{L}$, I'm not sure.

3.4 The set of terms of the language is a term-algebra, with the function-symbols as its operations. The algebra is generated by the simple names, and variables. The universe of the interpretation, together with the functions in the domain of the interpretation function ι form an algebra of the same type. The object assignments are are the homomorphisms from the term algebra to this "object" algebra. These sets of maps are algebras of some sort too aren't they?

3.5 The language itself is also a term algebra, generated by the atomic sentences, truth-value-names ($\top_i, \bot_i$) and propositional variables. The matrix is an algebra of the same type, on the truth values. The truth-valuations are the homomorphisms from the language to the truth values.

3.6 I'm interested in the difference between the universal algebra type account of compound functions (which could probably do my kind of relations too) and the way implicit in the variables-and-assignments approach I'm on about. The algebraic story in Burris and Sankapanivar has holes in it [B&S00], though I can't remember what they were now. Function composition generating an algebra of functions: both truth and object functions.

3.7 Another similar approach is making an augmented language like the "contexts" of λ calculus. You could do stuff like prove whether or not every relation on a universe is expressible in a given language.

3.8

3.9 I have rather cheekily called the denotations of quantifiers *quantities*, mostly because it sounds nice. Certainly we can capture the ideas "all" and "some" with these notions. Could we also get stuff like 1, 2, π, i, $\aleph_0$, etc? Probably, but it would be kind of circular, like the use of the word "every" in the possibity constraint on interpretations the universal quantifier.

4. Deduction

4.1 I would like to give proofs of the same thing, say

$$(\forall x)(Fx \supset Gx), (\exists x)Fx \vdash (\exists x)Gx$$

, in each of the proof systems. Perhaps even a propositional example.

4.2 How do you prove $A, B \vdash A \wedge B$ in a Hilbert system?

4.3 I'd like to say a little more about the biography of Gentzen. He was another one who did heaps and died young.

4.4 Is there any useful sense in which formal proofs are sentences in the language recognised by the deductive system? I think I recall somebody saying that sequent calculi can be seen this way.

5. Consequence

5.1 *Compactness.* There are notions of compactness in topology (and hence analysis etc.), lattice theory ([B&S00, Defn. 4.13], [Bir40, p. 186]) and in logic ([End72, p. 136], [Hod97, Ch. 5, esp. §5.3 p. 138], [Van71, p. 36], [Doe96, §4.1]). The lattice version is just a generalisation of that of topology, but compactness in Logic seems a bit different. Only a bit though. I expect they are somehow dual, and that clarifying the Galios correspondence (see [Bir40]) in the logic will help to sort it out. The exercise suggested in [End72] is probably a good starting point.

5.2 *Consequence and topology.* A topological closure operator is a special kind of closure operator. Most of the consequence relations we study have *theorems* - sentences that are consequences of the empty set. That is $Cn(\emptyset) \neq \emptyset$, so consequence does not usually give you a topology over the language. One case where it does is in Kleene's K_3 (see [Pri01, §7.3], [Kle52, §64], [Kle38]). I'm also going to have a look more carefully at Vickers' *Topology via Logic* [Vic89]. Another tenuous connection: Combinatory logic is an alternative formulation of the lambda calculus, which in turn has a topological denotational semantics due to Scott. See [Bar84].

5.3 The consequence operator C induced by the consequence relation $\vdash$ over X is a topological closure iff $\vdash$ satisfies what?

6. Examples

6.1 I would liked to have discussed Lindenbaum algebras/matrices. If a set of sentences S is closed under substitution of variables, then you can factor the language by relating A and B iff $A \equiv B \in S$. In a way, this enables you to get a logic out of a theory. That is, just by knowing what is the case, you can work out what is a consequence of what. This seems well worth exploring further. a brief summary of my work ([OKe00]) on this kind of stuff would have been nice. That could also help with the Whitehead problem.

7. Non-standard Analysis

7.1 Enderton [End72, §2.8, p. 164] gives an account of Nonstandard Analysis that is highly compatable with my approach. Unfortunately, I discovered it too late. He uses an uncountable language.

7.2 I'll repeat here a call I made in a footnote: There ought to be a big catalogue of Logical principles, philosophical arguments for and against them, and logics that satisfy and fail them.

7.3 Must TV be a Boolean algebra to satisfy requirement 7.4? (page 45). Complements are not unique in nondistributive lattices. Consider the diamond lattice: each of the three middle elements is the complement of the other two. I'm not sure whether an ultrafilter-like thing could be defined on nondistributive lattices, that would allow us to satisfy the requirement. The proof of 7.5.8 depends on DeMorgan's laws. Do they depend on distributivity? There is probably a way of beginning with $\vee$ and ideals rather than $\wedge$ and filters, and if you did that, you'd need a Boolean Algebra to get 7.1.

7.4 The *transfer principle* definitely should be in here.

7.5 I had hoped to prove Lös' ultraproduct theorem for my generalised models. B. J. Gardner's "How to Make Many-Sorted Algebras One-Sorted" [Gar89] and Tarski's result that varieties are closed under homomorphisms, subalgebras and products would hopefully make the job easier. It may also be worth comparing Gardners' paper with Enderton [End72, §4.3]

7.6 A thought on substitution of equals: If φ is a formula with x free, let $\varphi(a)$ be φ with a substituted for x. We want $\models \varphi(b)$ whenever $\models \varphi(a)$ and $\models a = b$. Does making this so turn the universe into some kind of space where $\{\overline{va}(n) :\models \varphi(n)$ is an open set, and $\overline{va}(b)$ is a limit-point (or some such) of $\overline{va}(a)$? Is this yet another Galois correspondence (see [Bir40])? Compare with Zariski Topology (see [Har77]).

8. Mortensen [Mor00] thinks *functionality* (his term) rather than consistency is the key property that interesting mathematical theories must have. Functionality is substitutability of equals in logic free contexts. It is a weaker condition that transparency: substitutability of equals in all contexts. Does my framework permit non-functional theories, non-substitutive theories? This is related to my concerns in a footnote somewhere in Chapter 7.

9. My reading has been pretty deficient. I have not yet read [Pri97], [Pri96], [Pri87], or [daC74].

Bibliography

[Bir40] Garrett Birkhoff, *Lattice Theory*, American Mathematical Society, 1940 (1984)

[Bar84] H. P. Barendregt, *The Lambda Calculus: Its Syntax and Semantics*, Elsevier, 1984

[Bre91] Gerhard Brewka, *Nonmonotonic Reasoning: Logical Foundations of Commonsense*, Cambridge University Press, 1991

[B&S00] Stanley Burris and H.P. Sankappanavar, *A Course in Universal Algebra*, 2000 Millenium Edition, corrected version of the 1981 Springer edition, `http://www.thoralf.uwaterloo.ca/htdocs/ualg.html`

[B&R00] JC Beall and Greg Restall, Logical Pluralism, *Australian Journal of Philosophy*, 78 (2000) 475-493 `http://www.phil.mq.edu.au/staff/grestall/files/pluralism.ps`

[Cho56] Noam Chomsky, Three Models for the description of Language, *I.R.E* [Institute of Radio Engineers] *Transactions on Information Theory*, IT-2, pp.113-24

[Cho57] Noam Chomsky, *Syntactic Structures*, Gravenhage: Mouton, 1957

[Cur63] Haskell B. Curry, *Foundations of Mathematical Logic*, McGraw-Hill, 1963

[daC74] Newton C. A. da Costa, On the Theory of Inconsistent Formal Systems, *Notre Dame Journal of Formal Logic*, Volume XV, Number 4, October 1974

[Doe96] Kees Doets, *Basic Model Theory*, CSLI Publications & folli, 1996

[Ekl76] P.C. Eklof, Whitehead's Problem is Undecidable, *American Mathematical Monthly*, December 1976

[End72] Herbert B. Enderton, *A Mathematical Introduction to Logic*, Academic Press, 1972

[FS&S97] Desmond Fearnley-Sander and Tim Stokes, Equality Algebras, *Bull. Austral. Math. Soc.* Vol 56 1997

[F&M92] Harvey Friedman and Robert K. Meyer, Whither Relevant Arithmetic?, *The Journal of Symbolic Logic* Vol 57 Number 3, Sept. 1992

[Gar89] B. J. Gardner How to Make Many-Sorted Algebras One-Sorted, *Comment. Math. Univ, Carolinae* 30,4, 1989 pp. 627-635

[Göd31] Kurt Gödel, *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*, Dover 1992, reprinted from a translation of: Über Formal Unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, *Monatshefte für Mathematik und Physik*, v. 38 pp. 173-198, 1931

[Har77] Robin Hartshorne, *Algebraic Geometry*, Springer, 1977

[Hod97] Wilfrid Hodges, *A Shorter Model Theory*, Cambridge University Press, 1997

[How91] John M. Howie, *Automata and Languages*, Oxford University Press, 1991

[Hun00] Paul Hunter, *Ultra-Filters and Internalised Equality*, unpublished manuscript, May 2000

[H&L85] Albert E. Hurd and Peter A. Loeb, *An Introduction to Nonstandard Real Analysis*, Academic Press, 1985

[Jef91] Richard Jeffrey, *Formal Logic: It's Scope and Limits*, McGraw-Hill, 1991 (1967, 1981)

[Kei86] J. Keisler, *Elementary Calculus, An Infinitesimal Approach*, Prindle, Weber & Schmidt, 1986

[Kle38] Stephen C. Kleene, On Notation for Ordinal Numbers, *Journal of Symbolic Logic*, Vol. 3 pp. 150-155, 1938

[Kle52] S.C. Kleene, *Introduction to Metamathematics*, North Holland Publishing Company, 1952

[Kri80] Saul Kripke, *Naming and Necessity*, Blackwell, 1980

[Lem65] E. J. Lemmon, *Beginning Logic*, Thomas Nelson and Sons, 1965 (1987 and 1998 editions also exist)

[M&M84] Robert K. Meyer and Chris Mortensen, Inconsistent Models for Relevant Arithmetics *Journal of Symbolic Logic*, Volume 49, Number 3, September 1984

[Mor95] Chris Mortensen, *Inconsistent Mathematics*, Kluwer Academic Publishers, 1995

[Mor00] Chris Mortensen, Prospects for Inconsistency, in *Frontiers of Paraconsistent Logic*, Diderik Batens, Chris Mortensen, Graham Priest, P Van Bendegem (editors), Research Studies Press, 2000

[Nau63] P. Naur (ed.), Revised Report on the Algorithmic Language Algol 60, *Communications of the ACM*, 6:1, 1963, pp. 1-17

[OKe00] Greg O'Keefe, *Algebra for Logic*, unpublished manuscript, submitted for assessment in the unit "Deviant Logic" at the School of Philosophy, University of Tasmania

[Pri87] Graham Priest, *In Contradiction: A Study of the Transconsistent*, Martinus Nijhoff Publishers, 1987

[Pri97] Graham Priest, Inconsistent Models of Arithmetic Part I: Finite Models *Journal of Philosophical Logic*, 26 (2):223-235, April 1997

[Pri96] Graham Priest, *Inconsistent Models of Arithmetic Part II: the General Case* manuscript (I'm not sure if it has been or will be published)

[Pri01] Graham Priest, *An Introduction to Non-Classical Logic: Moving about in Worlds Not Realised* Cambridge University Press, 2001 (references here are to a pre-publication manuscript, and may therefore differ in the printed version)

[Qui61] W.V. Quine, Two Dogmas of Empiricism, in *From a Logical Point of View*, 2nd ed., Harvard University Press. Reprinted in A.P. Martinich, *The Philosophy of Language*, Oxford University Press, 1996

[R&B80] Nicholas Rescher and Robert Brandom, *The Logic of Inconsistency*, Oxford, 1980

[Res00] Greg Restall, *An Introduction to Substructural Logics*, Routledge, 2000

[Rob61] Abraham Robinson, Non-standard Analysis, *Proceedings of the Royal Academy of Sciences, Amsterdam*, ser. A, 64, pp. 432-440

[Rob74] Abraham Robinson, *Non-standard Analysis*, North-Holland, 1974. Reprinted 1996, Princeton University Press

[Rus18] Bertrand Russell, The Philosophy of Logical Atomism, in *The Monist*, 1918, reprinted in [Rus56]

[Rus24] Bertrand Russell, Logical Atomism, 1924, printed in [Rus56]

[Rus56] Bertrand Russell, *Logic and Knowledge*, Unwin Hyman, 1956 (1988)

[She74] S. Shelah, Infinite Abelian Groups - Whitehead Problem and Some Constructions, *Is. J. Math.*, 18, 1974

[Smu68] R. Smullyan, *First Order Logic*, Springer Verlag, 1968

[Sto74] Robert R. Stoll, *Sets, Logic and Axiomatic Theories*, second edition, W.H. Freeman and Company, 1974 (1961)

[Tar30] Alfred Tarski, *Fundamental Concepts of the Methodology of the Deductive Sciences*, Reprinted in [Tar56]

[Tar36] Alfred Tarski, *On the Concept of Logical Consequence*, Reprinted in [Tar56]

[Tar56] Alfred Tarski, (translated by J.H. Woodger) *Logic, Semantics, Metamathematics*, Oxford at the Clarendon Press, 1956

[Var95] Achele Varzi, Model-Theoretic Conventionalism, in James Hill and Petr Kottko (eds.), *Karlovy Vary Studies in Reference and Meaning*, Prague: Filosofia, 1995, pp. 406-430. `http://www.columbia.edu/ av72/papers/KarlovyVary_1995.pdf`

Kluwer Academic Publishers, 1999

[Var99] Achele Varzi, *An Essay in Universal Semantics*, Kluwer Academic Publishers, 1999

[Var01] Achele Varzi, *Pluralism, Conventionalism, and the Bounds of Logic*, Australasian Association of Philosophy Conference, Hobart (Australia), July 4, 2001.

[Van71] Bas C. van Fraassen, *Formal Semantics and Logic*, Macmillan, 1971

[Vic89] Steven Vickers, *Topology via Logic*, Cambridge University Press, 1989

[Wes96] Jane Kirchner West, A $\diamond$-Characterization of Some Small Totally Projective Groups, in *Abelian Groups and Modules: Proceedings of the International Conference at Colorado Springs*, edited by David M. Arnold and Kulamani M. Rangaswamy, Marcel Dekker Inc., 1996

[Wit22] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, Kegan Paul, 1922

[W&R10] A. N. Whitehead and B. Russell, *Principia Mathematica*, 3 vols, Cambridge University Press, 1910-1913, 2nd ed., 1925-1927

[Won98] Paul Wong, *Paraconsistent Inference and Preservation*, Technical Report, Automated Reasoning Group, ANU, `http://arp.anu.edu.au/ftp/techreports/1998/TR-ARP-11-98.ps.gz`

[Zal88] Edward N. Zalta, *Intensional Logic and the Metaphysics of Intentionality*, The MIT Press/Bradford Books, 1988