ON SOLOVAY'S THEOREM AND A PROOF OF ARITHMETICAL COMPLETENESS OF A NEW PREDICATE MODAL LOGIC

YUNGE HAO

A THESIS SUBMITTED TO THE FACULTY OF
GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING
AND COMPUTER SCIENCE
YORK UNIVERSITY
TORONTO, ONTARIO

April 2021

Abstract

This thesis investigates a first-order extension of GL called ML^3. We briefly discuss the latter's properties and some of its toolbox: some metatheorems, the *conservation theorem*, and its semantic completeness (with respect to finite reverse well-founded Kripke models).

Applying the Solovay technique to those models the thesis establishes its main result, namely, that ML^3 is arithmetically complete. As expanded below, ML^3 is a first-order modal logic that along with its built-in ability to simulate general classical first-order provability — *having* "$\Box$" *simulate* the *metamathematical* classical "$\vdash$"— is also arithmetically complete in the Solovay sense.

We also carefully reconstruct the proof of Solovay's Lemmata in our Appendixes, including a complete mathematically rigorous construction of his graph-walking function "h".

To Chenyan

Acknowledgements

I would like to thank my parents, who fully support me on my decision of graduate study. I offer my gratitude to my supervisor and mentor, Dr. George Tourlakis, who taught me logic and computability, and helped me a lot in my study, without whom I could never finish this thesis; also to my co-supervisor, Dr. Jeffery Edmonds, who first introduced to me the fascinating theory of computation, and provided me with the opportunity to work with professor Tourlakis; to Dr. Franck van Breugel for his advice, his warm daily greetings, and letting me use a nice working place in his lab. I thank all who helped me on my way here.

Table of Contents

Chapter 1

Preliminaries

This thesis is about a specific modal predicate logic ML^3 [31, 32] and its "*arithmetical completeness*". In this introductory section we ensure that terms such as the quoted above are defined here (to the best approximation that the limited resources of this section can provide) and that we present a brief outline of the various uses of *modal logic*, as such uses are accommodated by various *interpretations* of what the so-called "*modal box*" $\Box$ affords.

[Most of the definitions and terms of this chapter are adapted from the cited sources.]

1.1. Background and History [1, 26, 34, 38, 39]

The reader will be most likely familiar with the term "countable set A" as applied in the naïve (synonym of "informal", which means not based on axioms [not axiomatic]) set theory of Cantor. A is the range[1] of an everywhere defined function that has inputs from $\mathbb{N} = \{0, 1, 2, \ldots\}$ —the set of natural numbers.

A *theory* or *applied logic* is the predicate calculus (or first-order *classical* logic), based on axioms, equipped with additional, *mathematical* axioms that are pertinent to the theory this applied logic describes. If a first-order logic has no mathematical axioms, then it is a *pure* logic.

Examples of applied logics are Peano Arithmetic and ZFC axiomatic set theory. The acronym is derived from **Z**ermelo **F**raenkel

[1] Some terms such as "range" are covered in an elementary discrete mathematics course and are not defined in this thesis.

axiomatic set theory, *with* the axiom of **C**hoice.

Regarding the italicised term "classical" in the preceding paragraph, it is applied in a context such as ours in this thesis where we also discuss *modal* logics, that is, logics that have the symbol $\Box$ (called "Box" or "modal Box") among their *logical connectives* (that is, $\Box$ is additional to the *classical* connectives $\neg, \wedge$). Thus a *classical* logic (whether Boolean or first order) has no $\Box$ among its connectives; otherwise it is a *modal* logic.

A *theory* is a device *we* construct with a view of doing mathematics *using it as a tool* to prove theorems,[2] the idea being (originating with Hilbert) that 1) doing mathematics in a precisely formulated *formal* (based on *form*: syntactic) logic we avoid pitfalls of careless informal (not based on axioms and syntax) semantic reasoning where belief of knowledge is more dangerous than lack of it. 2) This is rather tenuous: Hilbert believed that a formal theory is necessarily amenable to a (by finite means, finitary, without delving into infinite sets) proof of its consistency (freedom from contradiction).

However this was dispelled by Gödel's *second incompleteness theorem* that said Peano Arithmetic (PA) *cannot prove its own consistency*. The relationship with 2) is this, in outline: If we *could* prove the consistency of PA with finite means, said finite means/processes could be embedded[3] and reasoned in PA, contradicting the second incompleteness theorem.

Reason 1) found strong proponents in Dijkstra and Scholten [11] and subsequently in Gries and Schneider [17], an undergraduate discrete mathematics text based on formal logic.

Incidentally, the consistency of PA *is* proved *outside* the theory PA easily by constructing a model (this construction is *not* finitary!) of the theory. As this model construction is outside the theory we say that we build and use it in the *metatheory* that often goes by the nickname "*real mathematics*".

Using a theory to *reason* (syntactically) we are *arguing formally*. When arguing within "real mathematics" then we are arguing *informally* (in the metatheory).

What *is* a model of a theory?

Consider a toolbox of two ingredients: A set $D \neq \emptyset$ (often called

[2]Much like building scissors to use them to cut paper, cloth and other things.

[3]By arithmetisation, a technique invented by Gödel in the proof of his first incompleteness theorem.

the domain of interpretation) and an *interpretation mapping*, M, that maps abstract symbols and strings —all the way up to formulae— of the theory to concrete ones that use the members of D as building blocks. We call the pair D and M collectively *a structure* $\mathscr{D} = (D, M)$ *for* our theory. By choosing different D and M we have in principle infinitely many choices for $\mathscr{D}$ if our theory has an infinite alphabet.

We build such a structure for ML[3] in detail in Chapter 6.

If an abstract formula A is interpreted into a concrete formula A', then we write this as $M(A) = A'$ or preferably (in the literature) $A' = A^{\mathscr{M}}$.

Now if A' has no free variables (in such an interpretation it will not; cf. [39]), that is, it is a *sentence*,[4] then we write $A^{\mathscr{M}} = \mathbf{t}$ *if we find that the concrete* A' *is true* ($\mathbf{t}$ for true and $\mathbf{f}$ for false), but preferably we write

$$\models_{\mathscr{M}} A$$

We say that A (note: not A'!) is true in the structure $\mathscr{M}$, or better still, $\mathscr{M}$ is a model of A.

We say that $\mathscr{M}$ is a model of a theory if it is a model of *each* of the axioms of the theory, those it has by being a *logic*, and those it has in order to describe the mathematics of the theory (mathematical axioms).

We write $\models A$ meaning "for all $\mathscr{M}$, $\models_{\mathscr{M}} A$". We say A is *valid*.

Gödel's *completeness* theorem says that in a pure first-order classical logic, all valid formulae are provable (that is, theorems).

Soundness is the easy converse: Every *pure* classical predicate calculus is sound, that is, if A is a theorem, then it is valid.

The above is *semantic* completeness. *Simple completeness* is a concept we define for PA and recursive extensions thereof (extensions by adding a recursive[5] set of axioms).

Such a theory —if consistent (free of contradiction)— is *simply complete* means that for every sentence A, the theory proves one of A or $\neg A$.

Gödel's first incompleteness theorem says that PA is simply *incomplete*: There is a sentence A, such that PA can prove neither A nor $\neg A$ (turns out there are infinitely many such A).

A corollary is that there are *true* sentences in "the standard model", $\mathscr{N} = (\mathbb{N}, M)$, of PA that are not syntactically provable (not

[4] Also, "closed formula".

[5] Algorithmically recognisable.

theorems).

Incidentally *the standard model* (also called "*intended model*" —it is the structure constructed/defined *in real mathematics* that PA *aims to describe* axiomatically) interprets all the symbols of the formal theory PA *as expected* over the domain $\mathbb{N}$: 0 is interpreted as 0, the S function (successor) is interpreted as the function $x \mapsto x+1$, $+$ is interpreted as the concrete $+$, etc.

The description of structure and model above deviates from *Tarski semantics* (see [34, 38], for example; this does not build a concrete formula interpretation) and is given in detail in [39]. Its advantage is that it facilitates understanding the definition of "*arithmetical interpretation*" and "*arithmetical completeness*" given below.

Sometimes a theory (especially in works on model theory) is given as the *set of all its theorems*. This the "extent of the theory" and we say the theory is given *extensionally*. This is how QGL_1 is given on p.10.

Other times it is given as earlier described: *Axiomatically*, capturing in a concise description the *tools* (axioms and rules of inference) that we may use repeatedly to *obtain* —in principle— all its theorems. We say it is given *intensionally*.

Arithmetical interpretation or *realisation* applies to *modal logics*, Boolean (like the logic GL of Solovay) or predicate logics like M^3, ML^3, QGL^b, QGL_2. It is a *syntactic interpretation*, imitating the M that we introduced earlier (in very rough outline) to map *abstract* formulae to *concrete* formulae of the metatheory or of the "real mathematics".

In an *arithmetical interpretation* something entirely analogous takes place: We map formulae of a modal logic to *formulae of PA* —*not* to "concrete" formulae in real mathematics. The usual symbol one sees in the literature for such an arithmetical interpretation (or mapping) is "*". This is defined arbitrarily (that is, as we please) on atomic formulae of the modal logic and proceeds by induction on the formation of modal formulae.[6]

For example for all classical connectives $(\neg, \wedge, \forall)$ it "commutes with them" meaning, for example, that $(A \wedge B)^*$ is defined as $A^* \wedge B^*$ and $((\forall x)A)^*$ is defined as $(\forall x)A^*$. The Box is interesting: We define $(\Box A)^*$ as $\Theta(\ulcorner A^* \urcorner)$, where $\Theta(x)$ is *the Gödel provability predicate* that

[6]Every such initialisation on atomic formulae uniquely define * on the remaining formulae as we know from a standard theorem on inductive definition on ordered sets. The set of formulae of a modal logic is ordered, for example, by formula complexity (number of connectives in the formula).

says "x is the Gödel number of a PA theorem" and "$\ulcorner B \urcorner$" denotes the Gödel number of the PA formula B.

Thus, once the interpretation is defined we obtain from a modal formula A a formula A^* of PA. *Our interest now is syntactic*: *Not* if A^* is *true* but rather *whether it is a PA theorem!*

Correspondingly, *arithmetical completeness* for some modal logic L is the statement: *For every formula A of L, if* for all choices of interpretation * *—determined by how we define * on atomic formulae— A^* is a PA theorem, then A is an L theorem.*

Some history:

The early twentieth century is famous for its many "crises" in science, a tendency to overthrow the classical world view loomed around many great minds. In mathematics, David Hilbert proposed a program to establish the consistency of formal axiom systems by finitary means —that is without delving into the infinite, for example not assuming the availability of (mathematical existence) nor using infinite sets such as the set of *all* natural numbers. What about induction? One can do induction within segments of $\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}$ such as $\{0, 1, 2, 4, \ldots, n-1, n\}$. We can have natural numbers as big as we please in our arguments (such numbers are built by a finite process), but we avoid collecting *all* of them into a *set*!

Investigating Hilbert's proposal Kurt Gödel utilised a version of a famous paradox, namely, "the liar paradox" of Epimenides of Crete in the form "I am lying".[7] He transferred this to the PA with a statement that said "I am not a theorem". Gödel actually proved that this sentence can be expressed as a formula (with no free variables; a sentence of logic) of PA and concluded with his famous first incompleteness theorem: Neither "I am not a theorem" nor its negation are provable in PA, hence PA is (simply) incomplete!

This was already bad for Hilbert's program, but the final blow was the second incompleteness theorem that stated "PA cannot prove its consistency."

[7] Actually the original form of Epimenides's paradox is reputed to say instead "All Cretans are liars", which is not as drastic as the form Gödel adapted. It simply implies that at least one Cretan who is not a liar exists.

1.2. Modal Logic in General [13, 16, 22]

Complex as classical logic goes, a truth is a truth and a falsehood is a falsehood, while what is beyond these two end points is forever left untouched. When the concept of truth stops being purely static, classical logic ends and modal logic begins. Modality originated from the need to express the notion of "necessity" and "possibility". A statement may have different modes of truthfulness, which can in turn be incorporated in the form of the statement. For example, a statement could be true at sometime, somewhere, or under some circumstances, while false otherwise. In words, modal logic is the logic of conditions placed on truth. More generally speaking, such a condition can either be scientific or a "belief". Here is a table from [16] that gives some examples of modal systems, grouped by the interpretation of modality ($\Box$):

temporal logic/tense logic	henceforth, eventually, hitherto, previously, now, tomorrow, yesterday, since until,inevitably, finally, ultimately, endlessly, it will have been, it is being...
denotic logic	it is obligatory that it is forbidden that, it is permitted/unlawful that
epistemic logic	it is known to X that, it is common knowledge that
doxastic logic	it is believed that
dynamic logic	after the program finishes after the computation ends, after the action is done, the program enables, throughout the computation
geometric logic	it is locally the case that
alethic logic	it is possible/necessary that it is contingent that it is impossible that
metalogic	it is valid/satisfiable that, it is provable/consistent that

In addition, *provability logic* can be singled out from metalogic as

an independent class of modal logic that studies "it is provable that ...", i.e., *provability*, which is *chosen* as the *interpretation* of $\Box$, or modality, in my thesis. I wish to emphasise that "provability where" is a subsidiary important question. Since Gödel's time, modality as provability in Peano Arithmetic has been the norm for provability logics. Motivated by a question posed by [18], [42, 43] introduced the study of modality in a first-order logic as *classical provability*. This study converged, with ML^3 to the beginning of the study of first-order logic modality as formalised provability in Peano Arithmetic. A contribution to the latter kind of research is the present thesis.

I wish to note that the word *chosen* at the onset of the paragraph below the preceding table cannot be emphasised enough. Researchers *choose* the interpretation of $\Box$ they need for the purpose of their own area of research. As $\Box$ is very versatile, it can be interpreted to fit a very large variety of applications, while not imposing or implying any of those interpretations by itself alone (unlike $\wedge$ and $\neg$ whose meanings are fixed and mean only one thing). A little bit of pre-shadowing, in this thesis we will see two totally different interpretations of $\Box$ with two modal logics, in one of the cases we have the one interpretation, while in the other case we have both of the two interpretations in the same time, in fact, one interpretation stepping upon the other!

Enough of the teasing, real work now! Modal logics, like classical logic, can be built by two approaches. Axiomatic method is one way to build modal systems. It defines the language of a system, a basic set of axiom schemata which is a selected set of well-formed formulae, and a set of transformation rules known as the *rules of inference*. The formulae derived from the axioms and rules of inference are the *theorems* of the system. An axiomatic approach is *syntactical*. By contrast, a *semantical* approach studies the meanings of the formulae and may also use semantics to informally prove theorems, as Cantor did...

Two branches of semantics study for modal logic were developed in the twentieth century: algebraic semantics, which focuses on Boolean algebra, and relational semantics, using Kripke structures. In the latter approach the symbol $\Box$ is adapted for expressing universality and the derived symbol $\Diamond$ for existentiality.

Modal logic as a theory is widely applied in different areas. Apart from the vast research devoted to modal logic in philosophy and linguistics, computer scientists use modal logic to do formal analysis of the behavior of computer programs. In artificial intelligence, modal

logic is needed for reasoning about dynamic states of the agents. Most recently, the flourish of game theory provides modal logic with new opportunity to bring its analytical power into play.

Chapter 2

Introduction

This thesis contributes to the metatheory of the modal predicate logic ML^3, the latter first introduced and studied in [31, 32]. Using the Solovay tool [37] we prove here that said logic is arithmetically complete (a term that we will define shortly).

While it is a fact that our *application* of Solovay's theorem to demonstrate arithmetical completeness of ML^3 can directly use his Main Lemmata without reference to their *proofs*, we felt writing this thesis as an opportunity to retell the proof of Solovay's results in full, including the main prerequsites, the Lemmata, with complete mathematically rigorous proofs that leave no questions in the mind of the reader. Neither Solovay [37] —and to the best of our knowledge— nor the literature at large, such as a number of influential references that followed him ([4, 10, 23, 35]), do full justice to the proofs.

Solovay introduced in loc. cit. the propositional provability logic GL (Gödel-Łöb logic) and proved that it is *arithmetically complete*, meaning that any GL formula is a theorem of GL if all its *arithmetical interpretations* are provable in Peano Arithmetic (PA). This particular version of completeness gives GL the name *provability logic* since it models the behaviour of *provability in PA*.

There has been a lot of interest in discovering first-order provability logics (cf. [4]). The obvious idea seemed to be to define *extensionally* a "Quantified GL" —or QGL_1[8]— as the set of *theorems* P below, over a first-order *classical* language augmented by the modal $\Box$ that has the property that $\Box A$ has the same free variables as A for all

[8] There is a QGL_2 —our numbering meaning to distinguish the two— that we will simply call QGL.

formulae A. *We call such a* $\Box$ *"transparent".*

$$QGL_1 \stackrel{Def}{=} \{P : \vdash_{PA} f(P) \text{ for } \textit{every} \text{ arithmetical interpretation } f\} \quad (1)$$

In [4], Boolos —speaking with the authority of one of the foremost logicians of the twentieth century— wrote: "... one major question then (1979) open was whether *quantified* provability logic could be axiomatized." He described how excited he was when he received a letter from Vardanyan from the then Soviet Union claiming the question was solved, however, his excitement was not well placed for he found out that the answer to the question was *negative*.

It is clear that as *provability logics* go —that is, in Boolos's words, "when modal logic is applied to the study of provability, it becomes provability logic"— there is a keen interest in modal quantified logics applied to proof theory (of PA).

Vardanyan [44] showed that this first-order logic QGL_1 is not recursively axiomatisable; in fact he proved a stronger result: The QGL_1 of (1) is Π^0_2-complete.[9]Thus the idea of taking "QGL" extensionally failed badly as we cannot make it into a tangible axiomatic system that is *usable*.

Another "QGL" was built "forward" rather than "backward", namely, as an already (recursively) axiomatised first-order extension of GL over the same language as (1), with the same behaviour of $\Box$ vis a vis free variables.

This *intentional* logic —"QGL_2" that we will name in this work "QGL" without a subscript— being necessarily *different* from the QGL_1 above in view of Vardanyan's result, has a minimal (finite) set of modal axioms added on top of the usual first-order classical axioms (cf. [3, 29, 41]). It turns out that this QGL has shortcomings too:

- It has no cut-free *Gentzenisation*, i.e., no cut-free Gentzen-equivalent logic (cf. [3]).
- It is *not complete* with respect to *any* class of Kripke frames and it *is not arithmetically complete*, both of the last two negative results being due to Montagna [29].

[9]The term means that (1) QGL_1 can be expressed as $(\forall x)(\exists y)P$ for some recursive P, and (2) all arithmetical relations of the form $(\forall x)(\exists y)R$, for recursive R, are *strongly reducible* to the set QGL_1, in symbols $(\forall x)(\exists y)R \leq \mathrm{QGL}_1$. Note that for any sets $A \subseteq \mathbb{N}^n$ and $B \subseteq N$ we write $A \leq_m B$ —and say A is strongly reducible to B— iff for some recursive $f : \mathbb{N}^n \to \mathbb{N}$ we have $\vec{x} \in A$ iff $f(\vec{x}) \in B$.

In [2, 23] a first-order QGL_2-like extension of GL is investigated and proved to be arithmetically complete. The authors adopt a "transparent" $\Box$ as is the case with QGL_1 and QGL_2.

While they succeed proving arithmetical completeness (with * interpreting modal formulae into PA), on the other hand for this to work they postulated some overly restrictive requirements on their "finite" Kripke models: The *domains* of each world were required to be *finite* as well, and moreover they had to satisfy certain inclusion relations between themselves at that, thus denying broad applicability of such modal logics.

Later on, Yavorsky [45] modified QGL into QGL^b, where this time the modal operator $\Box$ binds all free variables in a formula making them invisible: every $\Box A$ is a sentence. *We may say that such a Box is "*opaque*".*

QGL^b is recursively axiomatised and its primary rules are modus ponens, *strong generalisation* $A \vdash (\forall x)A$ and *strong necessitation* $A \vdash \Box A$. He proved that it is arithmetically complete (with * interpreting modal formulae into *a finite extension of* PA).

A closely related first-order logic is the ML^3 of [32], with essentially the same set of axioms and also with an opaque $\Box$, but for technical reasons this predicate modal logic has only the first two of the above rules as primary, "hiding" necessitation in the axioms in the style of [35]. Thus, ML^3 has an *admissible* rule of *weak* necessitation instead: If $\vdash A$ then $\vdash \Box A$.

It has been long understood by the research community working on predicate modal logics of provability that the failure of the attempts to obtain a first-order recursively axiomatised provability logic was due to the insistence on having a transparent $\Box$.

Indeed, the concluding remarks in [45] note the detrimental effects of a transparent $\Box$ on arithmetical completeness. Thus, while the earlier paper of [2] obtained arithmetical completeness of a predicate modal logic *with* a transparent $\Box$ it did so severely restricting finite Kripke models as we noted above.

Yavorsky [45] successfully experimented with an *opaque* $\Box$ and with the restricted Barkan formula

$$\Box A \to \Box(\forall x)A \tag{2}$$

as one of his axioms and showed that QGL^b is a first-order provability logic. His paper does not explain the significance of the choice of (2) —see however [41–43] who chose this axiom for reasons *totally*

unrelated to arithmetical completeness.

Through a different route, with some interesting intermediate stops, [32, 42, 43] arrived at the logic ML^3 that is the logic of interest in this thesis, while [41] further explored the significance of axiom (2) in ML^3 and M^3, in particular proving

- (2) is independent from the other axioms.
- If removed, the resulting logics are arithmetically *incomplete*.

Thus, all other axioms being left *as is*, [41] establishes that (2) is *essential* for arithmetical completeness. [45] does not comment on this matter at all.

Two precursors of the ML^3 are M^3 and BM of [42, 43]. These are extensions of the propositional modal logic K4, and were introduced in [24, 42, 43], both to satisfy the *conservation requirement*, that is, essentially, for any classical formula A, A is provable classically iff $\Box A$ is provable modally in M^3 (and thus also in ML^3). More precisely, this requirement states, for any classical formulae A and B and classical theory $\mathcal{T}$, that $\mathcal{T} \cup \{A\}$ proves B classically iff $\mathcal{T}$ proves $\Box A \to \Box B$ modally in M^3. In BM that opts for only *one* primary rule of inference (modus ponens) and thus both generalisation and necessitation are admissible rules, one has that $\mathcal{T} \cup \{A\}$ proves B classically iff $\mathcal{T}$ proves $\Box(A \to B)$ modally in BM.

Thus, in M^3 and BM, and for classical A, "$\Box A$" means, essentially, "*A is a classical theorem*", that is, $\Box$ formalises the classical metatheoretical notation $\vdash$ when it is applied to classical formulae.

The introduction of these logics was to attain said "conservation" result thus to obtain a "provability logic", not for PA, but rather for *pure classical predicate logic*, especially useful in the practice of *equational proofs* of [11, 17, 39]. The goal —triggered by an observation in [18]— to have $\Box$ imitate $\vdash$ for classical logic led to two design criteria:

- $\Box$ in M^3 (and later ML^3) *has* to be opaque, that is, $\Box A$ is closed for all formulae A, since for classical first-order strong generalisation logic (cf. [28, 34, 38]) we have $A \vdash (\forall x)A$. Thus $\vdash$ is *oblivious* to any free variable x of A and *therefore so must be* the modal counterpart of $\vdash$, that is, $\Box$.
- We have strong generalisation in M^3 (and ML^3), that is $A \vdash (\forall x)A$, and thus we *must* adopt the axiom $\Box A \to \Box(\forall x)A$ that

"says" the same thing in the modal language. Incidentally, the emphasised "must" in this bullet was not known to the authors of [24, 32, 42, 43] but was established later in [41].

The above interpretation of $\Box$ in M^3 and ML^3 is totally different from the interpretation of the $\Box$ in GL. The box operator of GL is interpreted *arithmetically* as $\Theta(x)$. The Σ_1-formula[10] $\Theta(x)$ stands for $(\exists y)Proof(y,x)$ where $Proof(y,x)$ states that there is a Gödel number y that codes a PA-proof of the formula with Gödel number x, or simply, "x codes a theorem in PA".

Now, by induction on formulae, if A is interpreted as A^*, then $(\Box A)^*$ is interpreted as $\Theta(\ulcorner A^* \urcorner)$ —which says "A^*" is a PA-theorem— where $\ulcorner X \urcorner$ denotes the Gödel number of X [15, 35]. This *, called a *realisation* or *interpretation* of a modal logic in Peano Arithmetic, and is a mapping from formulae of a modal logic to formulae of PA. It is unique once we define it arbitrarily on atomic formulae of the modal logic.

The logic ML^3 was introduced in [32], adding Łöb's axiom $\Box(\Box A \to A) \to \Box A$ to M^3, [11] thus ML^3 is a first-order extension of both GL and M^3, and thus can (provably) simulate classical provability $\vdash$ through $\Box$ as well. ML^3 is over the same language as its predecessors M^3 and BM, and in particular, $\Box A$ is closed for all A. Notice that, in M^3, $\Box$ originally means "$\Box A$, with classical A, says: 'A is classically provable.'" With the introduction of ML^3 adding Löb's axiom, $\Box$ means two things: the above, plus, independently, $\Theta(\ulcorner A^* \urcorner)$.

Now, an interlude regarding Löb's axiom. Löb's axiom is a modal counterpart of Löb's Theorem in PA, which states that for any sentence A, $\vdash \Theta(\ulcorner A \urcorner) \to A$ implies $\vdash A$. If this is formalised in PA (i.e., if we have PA "state it" using "$\Theta(x)$") it becomes $\vdash \Theta(\ulcorner \Theta(\ulcorner A \urcorner) \to A \urcorner) \to \Theta(\ulcorner A \urcorner)$.

Translating back to the modal language we get indeed the modal axiom $\Box(\Box A \to A) \to \Box A$.

[12,31,32] developed the proof theory for M^3 and ML^3 by devising cut-free Gentzenisations of each, called GTKS and GLTS respectively. Using a Gentzen logic as a proxy to study the proof theory of some Hilbert-style logic is a well-known methodology that profits from the

[10] A formula equivalent to one of the form $(\exists y)Q(y,\vec{x})$ —with recursive $Q(y,\vec{x})$— also known as *semi-recursive* or even *recursively enumerable* (r.e.)

[11] While Łöb's axiom can prove the axiom $\Box A \to \Box\Box A$ of M^3 and BM, we will retain it here as was done in loc. cit. for technical convenience.

subformula property[12] of cut-free Gentzen proofs.

[33] introduced certain formula to formula mappings named *formulators* (*formu*la trans*lators*). Such mappings preserve proofs in M^3, ML^3, and QGL, that is, if $\Gamma \vdash A$ holds in any one of the three logics, then for any well-chosen formulator $\mathfrak{F}$, $\mathfrak{F}(\Gamma) \vdash \mathfrak{F}(A)$. The formulators tool allows one to do metamathematical investigations directly on Hilbert-style proofs without Gentzenisation, bypassing messy cut elimination proofs. Even for QGL, a logic that *provably* does *not* admit cut elimination ([3]), the formulators tool was applied profitably ([33, 41]).

[32] proved the completeness of ML^3 with respect to *finite* reverse well-founded Kripke models, and also its *arithmetical soundness*. Because of this, and looking back at Solovay's proof [37] which heavily hinges on such finite Kripke models, the authors conjectured the arithmetical completeness of ML^3 in the conclusions section of loc. cit. (cf. also the introduction section of [41]).

The present thesis proves this conjecture, adapting the idea from [45] to work with a finite consistent *extension* of PA rather than PA itself.

Thus ML^3 is a new example of a *predicate provability logic* that can also simulate equational classical proofs.

[12] Every subformula of the conclusion appears as such also at every step of the proof.

Chapter 3

Language and Symbols

In this thesis, it being work on the metatheory of ML^3, we will follow the lead of other similar metatheoretical studies of predicate model logics such as QGL and $\mathrm{QGL}^{(b)}$ ([2, 23, 29, 45]) in simplifying the language of the logic under study. The simplified ML^3 here —just as QGL and $\mathrm{QGL}^{(b)}$ in loc. cit.— will contain no equality symbol ("=") nor any function or constant symbols (but will contain predicate symbols).

We will not go over the well known inductive definition of formulae over a first order alphabet (cf. [34, 38, 42]), however we note that a new connective, "the box $\Box$" is added to the alphabet of the classical language also adding this inductive formation clause to the definition of formulae: "if A is a first-order modal formula, then so is $\Box A$".

In this connection we note that while, as a technical convenience towards effecting Gentzenisation [32], one often separates object variables *a priori* into *free* and *bound* types. However, in our application of Solovay's theorem in this thesis we find it simpler to follow the *standard* syntactic approach where bound vs. free is determined by how the variable is *used* syntactically without an *a priori* such separation of types.

Below we note the key notational conventions in this thesis.

We use specific bold lower case latin letters, with or without primes or subscripts, $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, $\mathbf{m}$, for arbitrary *imported* constants from $\mathbb{N}$ that we will need in the semantics section. Logical truth values *true* and *false* are denoted by $\mathbf{t}$ and $\mathbf{f}$, respectively.

Formulae are denoted by capital latin letters A, B, C (with or without primes or subscripts). The formal logical connectives are $\neg$,

$\wedge$, $\forall$, $\Box$.

$\forall A$ denotes *the universal closure* of A, that is $(\forall x_1)(\forall x_2)\ldots(\forall x_n)A$, where the *list* $x_1, x_2, \ldots, x_n$ includes *all* the free variables of A *in lexicographic order. Thus, if* A *has no free variables —when we say "it is* closed*" or "it is a* sentence*"— then "*$\forall A$*" is the same string as "*A*".*

A sequence of variables is sometimes also denoted by the vector notation $\vec{x}_n$ or just $\vec{x}$ if the n is understood or unimportant.

We call a formula $\Box A$ *boxed. It is always a sentence.* In [41] an elaborate definition gives the exact syntax of $\Box A$, rendering the string "$\Box A$" *metanotation.* Here we will simply say that "the metanotation $\Box A$ means $\Box\forall A$". A formula is a *classical* formula if it does not contain $\Box$, otherwise it is a *modal* formula.

The connectives $\vee$, $\rightarrow$, $\leftrightarrow$, $\exists$ are introduced *via* the usual *definitions.* To reduce brackets in informal writing we assume the usual connective priorities and that they are all *right-associative.* $\Leftrightarrow$ is metatheoretical *conjunctional* equivalence synonymous with "iff". That is, $A \Leftrightarrow B \Leftrightarrow C$ means "$A \leftrightarrow B$ *and* $B \leftrightarrow C$". Similarly, $\Rightarrow$ is the conjunctional counterpart of "$\rightarrow$", that is, $A \Rightarrow B \Rightarrow C$ means "$A \rightarrow B$ *and* $B \rightarrow C$".

Capital Greek letters (with or without primes or subscripts) that do not match a Latin capital letter, e.g., Γ, Δ, Φ, etc., denote *sets* of formulae. $\Box\Delta$ denotes $\{\Box A : A \in \Delta\}$, $\forall\Delta$ denotes $\{\forall A : A \in \Delta\}$, and $\Delta\Box$ denotes $\{\forall A : \Box A \in \Delta\}$.

We write $B[x := y]$, $B[z := \mathbf{i}]$ and $B[q := A]$ to denote substitution into targets x, z, q in B. $A(x, u, w)$ conveys that x, u, w are *all* the free variables of A while $A[x, u, w]$ conveys that *possibly*, x, u, w are *among* the free variables in A. In the former case we may write $A(\mathbf{i}, u, w)$ for $A[x := \mathbf{i}]$ and in the latter case we may write $A[\mathbf{i}, u, w]$ for $A[x := \mathbf{i}]$.

Chapter 4

The Logic ML^3

As already noted in the previous chapter, the language L of ML^3 in the present thesis will have predicate symbols but no function symbols or constants. However, the language will later be augmented (cf. 5.0.2 and 5.0.5) to include *imported* constants.

4.1. Axioms

DEFINITION 4.1.1 (Basic Axiom Schemata of ML^3).

A1 All tautologies

A2 $(\forall x)A \to A[y]$ and $(\forall x)A \to A[\mathbf{k}]$, if $\mathbf{k}$ is a *constant* (cf. 5.0.2 and 5.0.5 that refer to *imported constants*). The result $A[y]$ is undefined if "y is captured by a quantifier" as in, e.g., [38].

A3 $A \to (\forall x)A$, if x does not occur free in A

A4 $(\forall x)(A \to B) \to (\forall x)A \to (\forall x)B$

A5 $\Box(A \to B) \to \Box A \to \Box B$

A6 $\Box(\Box A \to A) \to \Box A$

A7 $\Box A \to \Box(\forall x)A$

A8 $\Box A \to \Box\Box A$

Everywhere above we assume the normal priorities of connectives, that is from highest to lowest as follows

$$\overbrace{\neg, \forall}^{\textit{equal, highest}}, \wedge, \vee, \to, \equiv$$

Moreover in a chain $\ldots A \circ A' \circ A'' \ldots$ we associate from right to left. Thus, $(A) \to (A') \to (A'')$ means $(A) \to ((A') \to (A''))$.

The set of *all instances* of the schemata A1 – A8 is denoted by Λ. The set of (closed) *axioms* is $\forall\Lambda \cup \Box\Lambda$. The inclusion of $\Box\Lambda$ is the "Smoryński trick" that "hides" weak necessitation in the axioms.

$\Box A \to \Box\Box A$ can be derived in ML^3 from the schema A6, but is included for convenience to avoid also adding $\Box\Box\Lambda$ to the axioms.

[32] has introduced and studied a variant of ML^3 above, *with* function and constant symbols and *with* equality (and its axioms) included. As we already noted, it is simpler —and customary ([2, 23, 29, 45])— to discuss arithmetical completeness without these features.

The inclusion of A7 is based on remarks in [18] and was added already in the sublogic M^3 introduced in [42] as a necessary axiom toward enabling $\Box$ to simulate the classical *metasymbol* $\vdash$.[13] $\Box$

4.2. Rules of Inference

Definition 4.2.1. The rules of inference of ML^3 are two, modus ponens (MP) $A, A \to B \vdash B$ and (strong) generalisation $A \vdash (\forall x)A$.[14]

$\Gamma \vdash A$ (resp. $\vdash_\Gamma A$) in ML^3 means that A is derived from axioms and hypotheses Γ (resp. hypotheses $\Gamma \cup \Box\Gamma$). Note that in a classical proof system $\vdash_\Gamma A$ means the same as $\Gamma \vdash A$. $\Box$

Unlike QGL^b where necessitation is postulated as a *strong* primary rule $A \vdash \Box A$, in ML^3 *weak* necessitation is admissible (cf. [32, 41, 42]).

Remark 4.2.2 (Tautological implication). One writes $A_1, A_2, \ldots, A_n \models_{taut} B$ pronounced "the $A_1, A_2, \ldots, A_n$ tautologically imply B". This means that $A_1 \to A_2 \to \ldots \to A_n \to B$ is a tautology, in symbols, $\models_{taut} A_1 \to A_2 \to \ldots \to A_n \to B$. $\Box$

[13] The presence of A7 as a *theorem* was trivially necessary, to simulate generalisation "$A \vdash (\forall x)A$" as "$\Box A \to \Box(\forall x)A$". That A7 *was not derivable* from the remaining axioms —and thus had to be an axiom— was first established in [41].

[14] This is equivalent to "$\Gamma \vdash A$ implies $\Gamma \vdash (\forall x)A$". *Weak* generalisation requires this Γ to contain no formula where x occurs free. This is a *global* restriction that makes the use of weak generalisation not so straightforward.

4.3. Metatheorems

Axiom group A1 immediately implies

THEOREM 4.3.1 (Proof by tautological implication). *If* $A_1, A_2, \ldots, A_n \models_{taut} B$*, then* $A_1, A_2, \ldots, A_n \vdash_{ML^3} B$*.*

THEOREM 4.3.2 (The Deduction (Meta)Theorem). *If* $\Gamma, A \vdash B$ *and* A *has no free variables (we say it is a* sentence*, or is* closed*), then* $\Gamma \vdash A \to B$

The above is a standard metatheorem and its proof is omitted. Cf. [28, 34, 38] for logics with a strong generalisation rule.

DEFINITION 4.3.3 (A consistent set of formulae (e.g., axioms)). A set of formulae Σ is *consistent* iff it fails to prove at least one formula.

Otherwise it is *inconsistent* (it proves all formulae).

THEOREM 4.3.4 (The inconsistency test). Σ *is inconsistent iff* $\Sigma \vdash \bot$.

PROOF:

- Σ is inconsistent. As it proves all formulae, in particular it proves $\bot$.
- Let $\Sigma \vdash \bot$. Pick any formula A. Now $\bot \models_{taut} A$ thus $\Sigma \vdash A$ by 4.3.1.

□

THEOREM 4.3.5 (Proof by Contradiction Metatheorem). *If* A *is a sentence, then* $\Sigma \vdash A$ *iff* $\Sigma, \neg A \vdash \bot$.

Note. We write Σ, B or $\Sigma + B$ for $\Sigma \cup \{B\}$, where B is a formula.

PROOF:

- Say $\Sigma, \neg A \vdash \bot$. By 4.3.2, $\Sigma \vdash \neg A \to \bot$. But $\neg A \to \bot \models_{taut} A$, hence $\Sigma \vdash A$ by 4.3.1.
- Let $\Sigma \vdash A$. Then $\Sigma, \neg A \vdash A$ by the definition of proof, and also, by the same definition, $\Sigma, \neg A \vdash \neg A$. Now, since $A, \neg A \models_{taut} \bot$, we get $\Sigma, \neg A \vdash \bot$ by 4.3.1.

□

THEOREM 4.3.6 (Weak Necessitation). *If* $\Gamma \vdash_{ML^3} A$, *where* $\Gamma = \Gamma' \cup \Box\Gamma'$ *or* $\Gamma = \Box\Gamma'$, *then* $\Gamma \vdash_{ML^3} \Box A$.

PROOF: (cf. [32, 41, 42]) By induction on $\Gamma-$theorems.

1. A is in $\Lambda \cup \Box\Lambda$.
   - $A \in \Lambda$. Thus $\Box A \in \Box\Lambda$ and hence $\Box A$ is an axiom. Done.
   - $A \in \Box\Lambda$. Then A is $\Box B$ with $B \in \Lambda$. By A8, $\Box B \to \Box\Box B$ is a theorem. By MP, so is $\Box\Box B$. But this is $\Box A$.
2. A is in Γ.
   - $\Gamma = \Gamma' \cup \Box\Gamma'$. If $A \in \Gamma'$ then $\Box A \in \Box\Gamma'$. Done. Otherwise, A is $\Box B$ with $B \in \Gamma'$. By A8 and MP we can prove $\Box\Box B$, that is, $\Box A$. Done.
   - $\Gamma = \Box\Gamma'$. See the second subcase of the previous bullet.
3. A is obtained by MP, that is, $\Gamma \vdash_{ML^3} B$ and $\Gamma \vdash_{ML^3} B \to A$. By I.H., $\Gamma \vdash_{ML^3} \Box B$ and $\Gamma \vdash_{ML^3} \Box(B \to A)$. By A5, $\Gamma \vdash_{ML^3} \Box(B \to A) \to (\Box B \to \Box A)$. Then we have $\Gamma \vdash_{ML^3} \Box B \to \Box A$ by MP. Combining with $\Gamma \vdash_{ML^3} \Box B$ and MP gives $\Gamma \vdash_{ML^3} \Box A$.
4. A is obtained by generalisation, that is, A is $(\forall x)B$ and $\Gamma \vdash_{ML^3} B$. By I.H., $\Gamma \vdash_{ML^3} \Box B$. By A7 $\Gamma \vdash_{ML^3} \Box B \to \Box(\forall x)B$, then by MP we have $\Gamma \vdash_{ML^3} \Box(\forall x)B$.

□

THEOREM 4.3.7 (The Conservation Theorem). *If* A *is a wff and* $\mathcal{T}$ *is a classical theory, then* $\vdash_{\mathcal{T}} \Box A$ *modally implies that* $\mathcal{T} \vdash A$ *classically (in* ML^3*).*

Chapter 5

Kripke Semantics

Kripke's possible worlds semantics [27] is the standard model theoretic approach to modal logic.

DEFINITION 5.0.1 (Kripke Frames). A *Kripke frame* is a pair $\mathcal{F} = \langle W, R\rangle$ where W is a non-empty set of *(possible) worlds* and R is a binary relation on W known as the *accessibility relation*.

We are interested in frames where R is transitive, irreflexive and *reverse well-founded* the latter meaning that there is no infinite R-chain $w'Rw''Rw'''\ldots$ □

DEFINITION 5.0.2 (Pointed Kripke Frames).

$\mathcal{F} = \langle W, R, w_0\rangle$ is a *pointed Kripke frame* if $\langle W, R\rangle$ is a Kripke frame and $w_0 \in W$ is a designated "initial" world. w_0 is selected to be *R-minimum*, called *the minimum node*, that is, $(\forall w \in W)(w = w_0 \vee w_0 R w)$. □

DEFINITION 5.0.3 (Primary Interpretation Mapping). Let L be a modal language, and let M_w be a non-empty countable set of objects, for each $w \in W$. I_w is an *interpretation* that *maps* the elements of L to the "concrete" domain M_w. It suffices to take each M_w to be enumerable since so is our alphabet and thus we take $M_w = \mathbb{N}$, for all $w \in W$. The I_w have the properties:

1. $I_w(q) \in \{\mathbf{t}, \mathbf{f}\}$ for every Boolean variable $q \in L$.
2. $I_w(\bot) = \mathbf{f}$ and $I_w(\top) = \mathbf{t}$.
3. $I_w(\phi) \subseteq \mathbb{N}^n$ for every predicate letter $\phi \in L$ of arity $n > 0$. □

We want a Henkin theory for L so rather than *assigning* (constant) values to variables we will *copy* values into variables. Values being

metalogical, the Henkin trick is to import them into the language L of our logic so that this copying becomes syntactically possible: Every $k \in M_w$ is imported as a formal constant $\mathbf{k}$. The resulting language is denoted by $L(M_w)$ ([34, 38]).

DEFINITION 5.0.4. If $A(x_1, \ldots, x_n)$ is over L, then we say that $A(\mathbf{k_1}, \ldots, \mathbf{k_n})$ over $L(\mathbb{N})$ is a sentence with *parameters* from M. □

The extended mapping for all closed formulae with parameters from M_w is defined as follows:

DEFINITION 5.0.5 (Extended Interpretation; forcing truth in a world.). By induction on *closed* formulae of $L(\mathbb{N})$, for every $w \in W$:

1. $I_w(\mathbf{k}) = k$, for each $\mathbf{k} \in L(\mathbb{N})$.
2. $I_w(\phi(\mathbf{k_1}, \ldots, \mathbf{k_n})) = \mathbf{t}$ iff $I_w(\phi)(k_1, \ldots, k_n) = \mathbf{t}$, for any n-ary predicate $\phi \in L$, where the k_i are in $\mathbb{N}$.
3. $I_w(\neg A) = \mathbf{t}$ iff $I_w(A) = \mathbf{f}$ for any closed formula A of $L(\mathbb{N})$.
4. $I_w(A \wedge B) = \mathbf{t}$ iff $I_w(A) = \mathbf{t}$ and $I_w(B) = \mathbf{t}$, for any closed formulae A and B of $L(\mathbb{N})$.
5. $I_w((\forall x)A) = \mathbf{t}$ iff $I_w(A(\mathbf{k})) = \mathbf{t}$, for all $k \in \mathbb{N}$, where $(\forall x)A$ is a sentence of $L(\mathbb{N})$.
6. $I_w(\Box A) = \mathbf{t}$ iff, for all w' such that wRw', we have $I_{w'}(\forall A) = \mathbf{t}$, where A is a formula of $L(\mathbb{N})$, closed or not.

If a sentence A over $L(M_w)$ satisfies $I_w(A) = \mathbf{t}$, then we write $w \Vdash A$. The notation $w \Vdash A$ is pronounced "w *forces* A". □

DEFINITION 5.0.6 (Kripke Structures). A Kripke structure for the modal language L is a pair $\mathcal{M} = (\mathcal{F}, \{(M_w, I_w) : w \in W\})$ where $\mathcal{F}$, M_w and I_w are defined as above. □

DEFINITION 5.0.7 (Truth in Kripke Models). For a modal language L and a modal formula A of L, a structure $\mathcal{M} = (\mathcal{F}, \{(M_w, I_w) : w \in W\})$ where $\mathcal{F} = (W, R, w_0)$ is a *Kripke model* of A, iff A is *true* in $\mathcal{M}$ at w_0, meaning $I_{w_0}(\forall A) = \mathbf{t}$, that is, $w_0 \Vdash \forall A$. *We can also write* $\models_{\mathcal{M}} A$ *in this case.*

We will not use the related concept of *validity* in a Kripke structure (defined as truth in *every* world) as it is equivalent to $w_0 \Vdash \Box A \wedge \forall A$.

For a modal language L and a set Γ of formulae of L, a structure $\mathcal{M}$ is a Kripke model of Γ iff $\mathcal{M}$ is a Kripke model of *every* A in Γ, written, *metatheoretically*, as $\models_{\mathcal{M}} \Gamma$.

Semantic implication of A from assumptions Γ, in symbols $\Gamma \models A$, means that every model of Γ is also a model of A; metatheoretically we may indicate this definition by "$(\forall \mathcal{M})\Big(\models_{\mathcal{M}} \Gamma$ implies $\models_{\mathcal{M}} A\Big)$". $\square$

Chapter 6

Semantic Completeness

This chapter proves the completeness of ML^3 with respect to *finite* Kripke models. It is based on the Kripke-completeness of M^3.

6.1. Semantic Soundness

The proof of *soundness* of ML^3 is sketched in [32]. It states,

PROPOSITION 6.1.1. For any given set of modal formulae Γ and any modal formula X, $\Gamma \vdash X$ implies that $\Gamma \models X$, and in particular, $\vdash_\Gamma X$ implies $\Gamma \cup \Box\Gamma \models X$, where semantics are over finite transitive and irreflexive Kripke structures.

For ease of reference we present a partial proof below: That every *modal* axiom is true in each world w. We omit the case of the truth of the classical axioms A1–A4 as it is is well covered in the literature ([28, 34, 38]).

Soundness then follows by the trivial observation that MP and generalisation preserve truth.

PROOF: Fix an $\mathcal{M} = (\mathcal{F}, \{(M_w, I_w) : w \in W\})$, and a world w.

1. X is A5, $\Box(A \to B) \to \Box A \to \Box B$. By the semantics of $\to$ and bracketing conventions (cf. p.17), let $I_w(\Box(A \to B)) = \mathbf{t}$ and $I_w(\Box A) = \mathbf{t}$. Thus *for each* w' *such that* wRw' we have

$$I_{w'}(\forall(A \to B)) = \mathbf{t} \text{ and } I_{w'}(\forall A) = \mathbf{t}$$

By the truth of A4 we obtain $I_{w'}(\forall B) = \mathbf{t}$.

2. The case where X is A6 is standard for GL (e.g., [35]) and is adapted to predicate logic.

   Suppose
   $$I_w(\Box(\Box A \to A) \to \Box A) = \mathbf{f} \tag{0}$$
   then
   $$I_w(\Box(\Box A \to A)) = \mathbf{t} \tag{1}$$
   and
   $$I_w(\Box A) = \mathbf{f} \tag{2}$$
   (2) entails that
   $$\text{for some } w', \text{ we have } wRw' \text{ and } I_{w'}(\forall A) = \mathbf{f} \tag{3}$$
   By (1), on the other hand, we have, for all w' such that wRw', that
   $$I_{w'}(\forall(\Box A \to A)) = \mathbf{t}$$
   and thus —via A4— for all w' such that wRw',
   $$I_{w'}(\Box A^{15} \to \forall A) = \mathbf{t} \tag{1'}$$
   Now —by (3) and (1′)— we conclude that
   $$I_{w'}(\Box A) = \mathbf{f} \tag{2'}$$
   and thus,
   $$\text{for some } w'' \text{ we have } w'Rw'' \text{ and } I_{w''}(\forall A) = \mathbf{f} \tag{3'}$$
   Now, (1′) is valid at w'' by transitivity of R. Thus we will have
   $$I_{w''}(\Box A \to \forall A) = \mathbf{t} \tag{1''}$$
   hence also
   $$\text{for some } w''' \text{ we have } w''Rw''' \text{ and } I_{w'''}(\forall A) = \mathbf{f} \tag{3''}$$
   and so on; we are generating an infinite chain
   $$wRw'Rw''Rw'''\ldots$$
   such that, for all n, $I_{w^{(n)}}$[16]$(\forall A) = \mathbf{f}$ and where the primed w are all distinct (by transitivity and irreflexivity). This contradict the reverse well-founded-ness of R thus (0) above is untenable.

[15] I should have written "$\forall\Box A$" here but, since $\Box A$ is *closed*, what I wrote is the same (equivalent).

[16] Where we wrote "$w^{(n)}$" to indicate w with n primes.

3. The case where X is A7 ([43]). If x is not actually free in A then there is nothing to prove. We need to show that $I_w(\Box A \rightarrow \Box(\forall x)A) = \mathbf{t}$. If $I_w(\Box A) = \mathbf{t}$, then for all w' such that wRw', $I_{w'}((\forall \vec{y})(\forall x)A) = \mathbf{t}$, where x and $\vec{y}$ are the free variables of A. By 5.0.5.6., $I_w(\Box(\forall x)A) = \mathbf{t}$.

4. The case where X is A8 ([43]). If $I_w(\Box A(\vec{x}_n)) = \mathbf{t}$,[17] then, for *all* w', wRw' implies that *all* $\vec{k}_n \in \mathbb{N}^n$ make $I_{w'}(A(\vec{\mathbf{k}}_n))$ true.

   Suppose now that $I_w(\Box\Box A) = \mathbf{f}$, then, for some w' we have wRw' *and* $I_{w'}(\Box A) = \mathbf{f}$).

   The later means that, for *some* w'' and $\vec{k}_n \in \mathbb{N}^n$, we have $w'Rw''$ —and hence wRw''— *and* $I_{w''}(A(\vec{\mathbf{k}}_n))$ false. A contradiction.

It is also easy to see that the two rules of inference preserve truth. □

6.2. Main Semantic Lemma

The Consistency Theorem [34, 35, 38] provides our first step towards proving the *Completeness* of ML^3 with respect to *finite* Kripke models.

The latter states $ML^3 \models A$ implies $ML^3 \vdash A$, where by "$ML^3 \models A$" we mean

$$(\forall \textit{ finite}, \text{ irreflexive, transitive } \mathcal{M}^f)(\models_{\mathcal{M}^f} ML^3 \text{ implies } \models_{\mathcal{M}^f} A) \qquad (\ddagger)$$

where the superscript f stands for "finite".

It turns out that we can obtain (‡) *from the proof of the Completeness of the subtheory* M^3 *via the latter's Consistency Theorem.*

THEOREM 6.2.1 (Consistency Theorem for a $\mathcal{T}$ over the language of M^3). *If a set of modal* sentences $\mathcal{T}$ *over the language of* M^3 *is consistent, then it has a Kripke model* $\mathcal{M}$.

PROOF: (cf. [43]) Recall that M^3 is ML^3 minus Löb's axiom.

Let $\mathcal{T}$ be a consistent closed modal *theory*.[18] For example, if we take $\mathcal{T}$ to be (intentionally) ML^3, then we prefer to let $\mathcal{T} = \forall\Lambda \cup \Box\Lambda$ —a set of *sentences*— which has the same deductive power as the

[17] The reader may want to review the notational conventions in Chapter 3.

[18] A closed theory *extensionally* is just a set of sentences; its closed theorems. *Intensionally* a closed theory usually is a set of rules and closed axioms intended to *generate* its set of theorems.

axiom set of ML^3 ($\Lambda \cup \Box\Lambda$) as it follows easily by axiom C.2.1(4.1.1) in one direction, and by strong generalisation in the other direction.

Let M be an arbitrarily selected *enumerable* set. *The same set is employed as the domain of individuals for each world.*

Firstly, we construct (loc. cit.) a *maximal* consistent extension of $\mathcal{T}$, called a *completion* of $\mathcal{T}$, following Henkin (for the classical case cf. [34, 38]). Since the language of M^3 is enumerable it is well-known that Henkin's method will work by taking $M_w = \mathbb{N}$, for all $w \in W$, for the sought Kripke model $\mathcal{M} = (\mathcal{F}, \{(M_w, I_w) : w \in W\})$. Of course, W, w_0 and R of $\mathcal{F} = \langle W, R, w_0 \rangle$ are yet to be determined.

Throughout we will take, for all w, $M_w = M = \mathbb{N}$

We fix enumerations of

1. all imported constants: $\mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{3}, \ldots$, bold face indicates the *formal* counterpart of $m \in \mathbb{N}$ in $L(M)$.

2. all closed modal formulae over $L(M)$: A_0, $A_1, \ldots$

3. all closed modal formulae over $L(M)$ of the form $(\exists x)A$: F_0, $F_1, \ldots$

   Note that each sentence $(\exists x)A$ of $L(M)$ is listed *infinitely often* (it is easy to see that this can be always arranged to be the case[19]).

Next we define by induction an infinite sequence of closed theories $\Gamma_0, \Gamma_1, \ldots$ by successively adding *sentences* over $L(M)$ to the sentences of $\mathcal{T}$ as new nonlogical axioms.

Basis step:

$$\Gamma_0 = \mathcal{T}$$

Induction step:

[19] For example, let $a_0, a_1, a_2, \ldots$ be a given enumeration. Form the infinite matrix $\begin{pmatrix} a_0 & a_1 & a_2 & \ldots \\ a_0 & a_1 & a_2 & \ldots \\ a_0 & a_1 & a_2 & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$ and enumerate its entries following the finite "south-east" diagonals.

Stage 1

$$\Delta_n = \begin{cases} \Gamma_n \cup \{A_n\} & \text{if } \Gamma_n \nvdash \neg A_n \\ \Gamma_n & \text{otherwise} \end{cases}$$

Stage 2

$$\Gamma_{n+1} = \begin{cases} \Delta_n \cup \{A[x := \mathbf{m}]\} & \text{if } \Delta_n \vdash F_{n+1} \text{ where } F_{n+1} \text{ is } (\exists x)A \\ \Delta_n & \text{otherwise} \end{cases}$$

where the bold face $\mathbf{m}$ is called a *Henkin constant* (or *witnessing constant*). It is chosen such that $\mathbf{m}$ is the *earliest* in the enumeration (1) (p.27) that does *not* occur in any of A_0, A_1, ..., A_n, F_0, F_1, ..., F_{n+1}.

We say that we added the sentence $A[x := \mathbf{m}]$ to Γ_{n+1} as a *special Henkin axiom.*

Now set $\Gamma = \bigcup_{n\geq 0} \Gamma_n$. This Γ satisfies the Main Semantic Lemma for M^3 listed below whose proof can be found in [43] and in AppendixA.

For any such completion Γ of $\mathcal{T}$, the central lemma is the following.

LEMMA 6.2.2 (Main Semantic Lemma for M^3, [34, 38, 43]).

Let $\mathcal{T}$ be a consistent set of modal sentences *over the language of M^3, and let M be an enumerable set (in our case $\mathbb{N}$). Then there is a* completion Γ *of $\mathcal{T}$ over $L(\mathbb{N})$ such that*

(1) $\mathcal{T} \subseteq \Gamma$

(2) Γ *is consistent.*

(3) Maximality. *For any sentence A over $L(\mathbb{N})$, either A or $\neg A$ is in Γ. This implies that Γ is* deductively closed, *i.e., $\Gamma \vdash A$ implies $A \in \Gamma$. The converse trivially holds.*

(4) Henkin Property. *If Γ proves the sentence $(\exists x)A$ over $L(\mathbb{N})$, then it also proves $A[x := \mathbf{m}]$ for some $m \in \mathbb{N}$.*

Now fix any *completion* Γ of $\mathcal{T}$ and call it w_0. Let Δ denote generically any completion of $\mathcal{T}$. We *define* (cf. [35, 43]) a relation R on the set of all such completions by

$$\Delta R \Delta' \text{ iff } \Delta\Box^{20} \subseteq \Delta'$$

[20] $\Delta\Box$ is defined in Chapter 3.

This R is transitive ([32, 35, 43]). To check this, suppose $\Delta R \Delta' R \Delta''$ and let $\forall A \in \Delta\Box$, we need to show that $A \in \Delta''$. Below we use the conjunctional "$\Rightarrow$" introduced in Chapter 3.

$$\begin{aligned}
\forall A \in \Delta\Box &\Rightarrow \Box A \in \Delta \ \langle \text{Definition of } \Delta\Box \rangle \\
&\Rightarrow \Box\Box \ A \in \Delta \ \langle \Delta \vdash \Box A \ 6.2.2(3) + \text{A8} + \text{MP} \rangle \\
&\Rightarrow \forall \Box A \in \Delta\Box \ \langle \text{Definition of } \Delta\Box \rangle \\
&\Rightarrow \Box A \in \Delta\Box \ \langle \Box A \text{ is closed} \rangle \\
&\Rightarrow \Box A \in \Delta' \ \langle \Delta R \Delta' \text{ and Definition of } R \rangle \\
&\Rightarrow \forall A \in \Delta'\Box \ \langle \text{Definition of } \Delta\Box \rangle \\
&\Rightarrow \forall A \in \Delta'' \ \langle \Delta' R \Delta'' \text{ and Definition of } R \rangle
\end{aligned}$$

Thus we let $W = \{w_0\} \cup \{w_a : w_0 R w_a\}$, discarding all inaccessible completions.

Define I_w as follows:

$$\text{For each Boolean variable } q \in L, I_w(q) = \mathbf{t} \text{ iff } q \in w \qquad (*)$$

For each n-ary predicate $\phi \in L$ and constants $k_i \in \mathbb{N}$,

$$I_w(\phi)(k_1, ..., k_n) = \mathbf{t} \text{ iff } \phi(\mathbf{k_1}, ..., \mathbf{k_n}) \in w$$

By 5.0.5, case 2,

$$I_w(\phi(\mathbf{k_1}, ..., \mathbf{k_n})) = \mathbf{t} \text{ iff } \phi(\mathbf{k_1}, ..., \mathbf{k_n}) \in w \qquad (**)$$

The next lemma is

For all modal sentences A over $L(\mathbb{N})$ and all $w \in W$, we have

$$w \Vdash A \text{ iff } A \in w \qquad (\dagger)$$

By $(\dagger)$ we are done with the Consistency Theorem: If $\mathcal{T}$ is consistent, then construct $\mathcal{M}$ as above. But then, if $\mathcal{T} \vdash A$ for some sentence A over L, then $w_0 \vdash A$ since $\mathcal{T} \subseteq w_0$. Thus $A \in w_0$ by deductive closure, hence $w_0 \Vdash A$ by $(\dagger)$. Thus $\mathcal{M}$ is a Kripke model of $\mathcal{T}$. $\Box$

Now we prove $(\dagger)$ by induction on the complexity of A.

PROOF: This is elaborated from [43], but as in [32] we define the complexity of $\forall B$ to be lower than that of $\Box B$.

1. A is atomic. This case is handled by $(*)$ and $(**)$.

2. A is $\neg B$.

   - If $w \Vdash \neg B$, then $w \nVdash B$ by 5.0.5, case 3. *Applying* I.H. and contraposition, $B \notin w$. By maximality, $\neg B \in w$.
   - If $\neg B \in w$, then $B \notin w$ by consistency. *Applying* I.H. and contraposition, $w \nVdash B$. By 5.0.5, case 3, $w \Vdash \neg B$.

3. A is $B \wedge C$.

   - If $w \Vdash B \wedge C$, by 5.0.5, case 4, $w \Vdash B$ and $w \Vdash C$. *Applying* I.H., $B \in w$ and $C \in w$. Hence $w \vdash B$ and $w \vdash C$. Then $w \vdash B \wedge C$. By deductive closure $B \wedge C \in w$.
   - If $B \wedge C \in w$, hence $w \vdash B \wedge C$, and thus $w \vdash B$ and $w \vdash C$. By deductive closure $B \in w$ and $C \in w$. *Applying* I.H., $w \Vdash B$ and $w \Vdash C$. By 5.0.5, case 4, $w \Vdash B \wedge C$.

4. A is $(\forall x)B$. If A has no free variables then we are done by the I.H. So let A be $A(x)$ —cf. Chapter 3.

   - Say, $w \Vdash (\forall x)B$, that is, $w \Vdash B(\mathbf{k})$, for all $k \in \mathbb{N}$. By the I.H. all the $B(\mathbf{k})$ are in w. Now if $(\forall x)B \notin w$ then the sentence $\neg(\forall x)B$ is in w by maximality of w; that is, $(\exists x)\neg B$ is. But then there is a Henkin witness $\mathbf{m}$ such that $\neg B(\mathbf{m})$ is in w, contradicting consistency.
   - Say, $(\forall x)B \in w$, hence $w \vdash (\forall x)B$. By axiom A2 and MP we have $w \vdash B(\mathbf{k})$, for *all* $k \in \mathbb{N}$. By deductive closure $B(\mathbf{k}) \in w$ —and by the I.H. $w \Vdash B(\mathbf{k})$— for *all* $k \in \mathbb{N}$. By 5.0.5, case 5, $w \Vdash (\forall x)B$.

5. A is $\Box B$.

   - Suppose $\Box B \in w$. Thus, using "$\Rightarrow$" conjunctionally (cf. Chapter 3)

$$\begin{aligned} \Box B \in w &\overset{\text{Def. of } R}{\Rightarrow} (\forall w')(wRw' \rightarrow \forall B \in w') \\ &\overset{I.H.}{\Rightarrow} (\forall w')(wRw' \rightarrow w' \Vdash \forall B) \\ &\overset{6 \text{ of } 5.0.5}{\Rightarrow} w \Vdash \Box B \end{aligned}$$

   - For the left-to-right part we proceed contrapositively. So let

$$\Box B \notin w \tag{1}$$

Let next $T = \{\neg\forall B\} \cup \{\forall C : \Box C \in w\}$. We write T as

$$T = \{\forall C_1, \forall C_2, \ldots, \neg\forall B\} \tag{2}$$

We claim that T *is consistent*. Proceeding by contradiction, *suppose otherwise*. Then (proof by contradiction) $\forall C_1, \forall C_2, \ldots \vdash_{M^3} \forall B$. Let $\{\forall C_{i_1}, \ldots, \forall C_{i_m}\}$ be the finite set of $\forall C_j$ used in this proof. Thus —by deduction theorem and $\forall B \to B$,

$$\vdash_{M^3} \forall C_{i_1} \to \ldots \to \forall C_{i_m} \to B$$

hence $\vdash_{M^3} \Box\forall C_{i_1} \to \ldots \to \Box\forall C_{i_m} \to \Box B$ by weak necessitation and A5 from 4.1.1. But all $\Box C_{i_j}$ —that is, $\Box\forall C_{i_j}$ by the definition of "$\Box X$"— are in w by definition of T, and this implies $\Box B \in w$ since w is deductively closed and contains the premises. We have just contradicted the main hypothesis (1) of this bullet.

Let then w' be a *completion* of the consistent T (3)

Now, $\forall B \notin w'$ since $\neg\forall B$ is in w' (consistency). By the I.H.,

$$w' \nVdash \forall B \tag{4}$$

If we can argue that we have

$$wRw' \tag{5}$$

then we are done since (4) and (5) imply $w \nVdash \Box B$. We do have (5′) by the definition of R on p.28.

□

6.3. Semantic Completeness

We next prove in detail that

THEOREM 6.3.1. *ML^3 is complete for finite, irreflexive and transitive Kripke models.*

We proceed contrapositively and start here:

$$\text{Assume for the sentence } A \text{ over } L \text{ that } \mathrm{ML}^3 \nvdash A. \qquad (\P)$$

By (¶), we have also $\mathrm{M}^3 \nvdash A$ since M^3 is a subtheory of ML^3. Thus by the preceding construction we have a Kripke model $\mathcal{M}$ for $\mathrm{M}^3 \cup \{\neg A\}$.

Using the "trick" of [32] below (6.3.5 and 6.3.7) we cut down the $\mathcal{M}$ model into a *finite, irreflexive, transitive* Kripke model, $\mathcal{M}^f$, of $\mathrm{M}^3 \cup \{\neg A\}$. As such $\mathcal{M}^f$ will be also *reverse well-founded* and hence also a model of ML^3 since it will satisfy also Löb's axiom. The details follow.

Remark 6.3.2. Note that every modal A can be put into a provably equivalent *normal form* where in each subformula of A of the form $\Box B$ the B can be replaced by $\forall B$. This is true —regardless of *the exact definition* of the "opaque" "$\Box$"— even if we did *not*, as we did, outright take "$\Box B$" to be a shorthand for "$\Box\forall B$". This is due to $\vdash_{M^3} \Box\forall B \leftrightarrow \Box B$ and the equivalence theorem.[21] Indeed, in one direction, note $\vdash_{M^3} \Box\forall B \rightarrow \Box B$ using repeated use of axiom A2, followed by weak necessitation and then repeated application of A5. In the other direction note $\vdash_{M^3} \Box B \rightarrow \Box\forall B$ by A7 followed by repeated application of axiom A5. □

"*Adequate sets*" of formulae occur in the literature in the construction of *finite* Kripke models and countermodels).

Definition 6.3.3 (Adequate set of formulae [23]). An *adequate set of formulae* Φ satisfies

1. It is *subformula-closed*, that is, if $A \in \Phi$, then all subformulae of A are also in Φ.
2. If $A \in \Phi$, then also $\neg A$ is in Φ where we apply recursively the rule of writing X for $\neg\neg X$. □

Definition 6.3.4. For any closed formula A in *normal form* — which without loss of generality has the form $\forall B$ for some B— over the language $L(\mathbb{N})$, the *augmented* set of subformulae of A, denoted by $S(A)$, is the *smallest adequate set that contains* A. Why "augmented"? Because the set of subformulae of A does not necessarily meet requirement 2 above.

[21] Replacing a subformula of a formula by a provably equivalent formula.

Note that not all formulae of $S(A)$ are closed. For example, if $(\forall x)B$ is a closed subformula of A, then B is in $S(A)$ but is *not* a *closed* subformula if $(\forall x)$ is not redundant. □

Trivially, $S(A)$ is a finite set. We next define a set of worlds W^f of the under construction finite Kripke structure and the related accessibility relation $\widehat{R}$. As in [32] we use the set $S(A)$ to help us "flag" the *members of the finite* subset W^f of worlds W that *we intend to keep*. Thus we define:

DEFINITION 6.3.5. Two worlds w and w' of the Kripke model $\mathcal{M}$ (for $\mathrm{M}^3 \cup \{\neg A\}$) above are said to be *equivalent*, in symbols $w \sim w'$, iff $w \cap S(A) = w' \cap S(A)$.[22] We take w_0 as the start world in W^f and we also select exactly *one* world from each *equivalence class* $[w]_\sim$ —where $w \nsim w_0$— to form a finite set of worlds W^f. Therefore the distinct worlds that we keep are the finitely many mutually non-equivalent worlds $w \in W$ as described.

To avoid *notational* confusion, if we selected $W^f = \{w_0, w_1, \ldots, w_{n-1}\}$ we *rename each such* w_i as α_i, so $W^f = \{\alpha_0, \alpha_1, \ldots, \alpha_{n-1}\}$.

(1) The accessibility relation $\widehat{R}$ on W^f is defined as follows

$\alpha \widehat{R} \beta$ iff both of the following bullets hold:

- For every subformula $\Box B$ of A, if $\Box B \in \alpha$, then $\{\Box B, \forall B\} \subseteq \beta$
- There is a subformula $\Box C$ of A in β such that $\Box C \notin \alpha$

(2) *Refine* W^f to omit redundant worlds: $W^f \overset{reset}{=} \left\{\beta : \alpha_0 \widehat{R} \beta\right\}$.

(3) Define $\alpha \Vdash_S F$ for all atomic *closed* $F \in S(A)$ to *mean* $F \in \alpha$. □

PROPOSITION 6.3.6. $\widehat{R}$ is reverse well-founded being (provably) irreflexive and transitive.

PROOF: We verify irreflexivity and transitivity. The consequence of this —reverse well-foundedness— is well known.

- *Irreflexivity.* Can we have $\beta \widehat{R} \beta$? If so, then some $\Box C$ (in $S(A)$) that is in the second copy of β will not be in the first copy of β (cf. second bullet of (1) above). Absurd.

[22] By its definition $\sim$ is trivially an equivalence relation.

- *Transitivity.* Let $\alpha \widehat{R} \beta \widehat{R} \gamma$. To prove $\alpha \widehat{R} \gamma$ let $\Box B$ be a subformula of A and $\Box B \in \alpha$. Then $\Box B$ (and $\forall B$) is in β. But then, by assumption, $\Box B$ and $\forall B$ is in γ. To conclude we check bullet two in condition of (1) above: Let the subformula $\Box C$ of A satisfy $\Box C \in \beta$ but $\Box C \notin \alpha$. But $\beta \widehat{R} \gamma$ implies $\Box C \in \gamma$. We are done.

□

LEMMA 6.3.7. *For A, α_i and $\widehat{R}$ as defined above and, for any* closed *X that is a subformula of A, we have $\alpha_i \Vdash_S X$ iff $X \in \alpha_i$.*

PROOF: This is essentially from [32] and is provided here for easy access. Induction on the complexity of X. Again, we define the complexity of $\forall B$ to be lower than that of $\Box B$.

1. X is an atomic sentence with parameters from $\mathbb{N}$. Done by Definition 6.3.5 (3).

   Two cases are more "interesting" than the others:

2. Case where X is $(\forall x)B$.

   - Say, $\alpha_i \Vdash_S (\forall x)B$, that is, $\alpha_i \Vdash_S B(\mathbf{k})$, for all $k \in \mathbb{N}$. By the I.H. all the $B(\mathbf{k})$ are in α_i. Now if $(\forall x)B \notin \alpha_i$ then the sentence $\neg(\forall x)B$ is in α_i by maximality of α_i; that is, $(\exists x)\neg B$ is. But then there is a Henkin witness $\mathbf{m}$ such that $\neg B(\mathbf{m})$ is in α_i contradicting consistency.
   - Say $(\forall x)B \in \alpha_i$, hence $\alpha_i \vdash (\forall x)B$. By axiom A2 and MP we have $\alpha_i \vdash B(\mathbf{k})$, for *all* $k \in \mathbb{N}$. By deductive closure $B(\mathbf{k}) \in \alpha_i$ —and by the I.H. $\alpha_i \Vdash_S B(\mathbf{k})$— for *all* $k \in \mathbb{N}$. By 5.0.5, case 5, $\alpha_i \Vdash_S (\forall x)B$.

3. Case where $X = \Box B$.

   - Suppose $\Box B \in \alpha_i$. It follows that

$$\begin{aligned} \Box B \in \alpha_i &\overset{6.3.5(1)}{\Rightarrow} (\forall \alpha_j)(\alpha_i \widehat{R} \alpha_j \rightarrow \forall B \in \alpha_j) \\ &\overset{I.H.}{\Rightarrow} (\forall \alpha_j)(\alpha_i \widehat{R} \alpha_j \rightarrow \alpha_j \Vdash_S \forall B) \\ &\overset{6 \text{ of } 5.0.5}{\Rightarrow} \alpha_i \Vdash_S \Box B \end{aligned}$$

- For the converse we proceed contrapositively. The proof with minor twists mirrors that of claim (†) on p.29.

  So let

  $$\Box B \notin \alpha_i \tag{1}$$

  Let next $T = \{\Box B, \neg\forall B\} \cup \{\Box C \in S(A) : \Box C \in \alpha_i\} \cup \{\forall C \in S(A) : \Box C \in \alpha_i\}$. We write T as

  $$T = \{\Box C_1, \forall C_1, \ldots, \Box C_m, \forall C_m, \Box B, \neg\forall B\} \tag{2}$$

  for some m. We claim that (this finite) T *is consistent*. Proceeding by contradiction, *suppose otherwise*. Then (proof by contradiction, followed by the deduction theorem) yields

  $$\Box C_1, \forall C_1, \ldots, \Box C_m, \forall C_m \vdash_{ML^3} \Box B \to \forall B$$

  Thus $\Box C_1, \forall C_1, \ldots, \Box C_m, \forall C_m \vdash_{ML^3} \Box B \to B$ (from $\forall B \to B$) hence $\Box C_1, \forall C_1, \ldots, \Box C_m, \forall C_m \vdash_{ML^3} \Box(\Box B \to B)$ by weak necessitation. Now by tautological implication (via Löb's axiom) we get $\Box C_1, \forall C_1, \ldots, \Box C_m, \forall C_m \vdash_{ML^3} \Box B$, which implies $\Box B \in \alpha_i$ since α_i is deductively closed and contains the premises. We have just contradicted the main hypothesis (1) of this bullet.

  Let then α_j be a *completion* of the consistent T (3)

  Now, $\forall B \notin \alpha_j$ since $\neg\forall B$ is in α_j (consistency). By the I.H.,

  $$\alpha_j \nVdash_S \forall B \tag{4}$$

  If we can argue that we have

  $$\alpha_i \widehat{R} \alpha_j \tag{5}$$

  then we are done since (4) and (5) imply $\alpha_i \nVdash_S \Box B$. So let $\Box C \in \alpha_i \cap S(A)$. Then $\Box C$ and $\forall C$ are in α_j (definition of T). Being subformulae of A we have established "half" of (5). For the other half we have (1) and also need that $\Box B \in \alpha_j$. This is true by (2) and (3).

  □

THEOREM 6.3.8. *ML^3 is complete with respect to* finite *reverse well-founded Kripke models (irreflexive and transitive).*

PROOF: To summarise, start at (¶), p.32. Then also $\mathrm{M}^3 \nvdash A$. Let $\mathcal{M} = (\langle W, R, w_0\rangle, \{(\mathbb{N}, I_w)\} : w \in W)$ be a model for M^3, where $w_0 \nVdash A$, as above. The model $\mathcal{M}^f$ for $\mathrm{M}^3 \cup \{\neg A\}$ on the frame $\langle W^f, \widehat{R}, \alpha_0\rangle$ constructed in the preceding discussion and used in 6.3.7 is a *finite* irreflexive and transitive model for M^3 hence also for ML^3 because of the implied reverse well-foundedness of $\widehat{R}$. Moreover we saw in 6.3.7 that $\alpha_0 \Vdash_S X$ iff $X \in \alpha_0$ for all $X \in S(A)$. In particular $\alpha_0 \Vdash_S A$ iff $A \in \alpha_0$, thus $\alpha_0 \nVdash_S A$ since $A \notin \alpha_0$ (recall, $\alpha_0 = w_0$ and also note (†) and its proof on p.29). □

Chapter 7

Arithmetical Completeness

7.1. Gödel Numbering and Peano Arithmetic [34, 38]

We have devised in the standard manner a model-theoretic approach to show the completeness of the logic system ML^3 with respect to finite Kripke models. In this chapter we turn to the matter of arithmetical completeness —to be defined shortly— of ML^3.

ML^3 is a *recursively axiomatized* theory, meaning it has a recursive set of axioms. Since the axioms are *strings*, we need to code them into *numbers, which we use* as proxies in order to speak of concepts from computability theory such as recursiveness and semi-recursiveness. A small "translation" or "coding" effort is needed in order to apply such concepts to string-theoretic functions and relations.[23] The coding-decoding algorithm is desired to be one-to-one and *intuitively computable*. In order to do that, we can use the prime power number sequence-coding scheme, invented by Gödel, who applied it in his investigation of the "simple completeness" of Peano Arithmetic (PA), this coding being named in the literature *Gödel numbering*. This scheme codes any *finite, ordered* sequence $a_0, a_1, \ldots, a_{n-1}$ from $\mathbb{N}$ as

[23] The term "number-theoretic" applied to functions and relations indicates that these objects take inputs from $\mathbb{N}$ and have outputs in $\mathbb{N}$ —the relations having their outputs in $\{0, 1\}$. Correspondingly "string-theoretic" functions and relations over some fixed alphabet take as inputs strings over said alphabet and produce strings as outputs.

the number

$$\prod_{i<n} p_i^{a_i+1}$$

denoted by $\langle a_0, a_1, \ldots, a_n\rangle$ and where p_i denotes the i-th prime ($p_0 = 2, p_1 = 3, p_2 = 5$, etc.)

Gödel assigned uniquely (in a one-to-one manner) *numbers* to each alphabet *symbol* of Peano arithmetic (see 7.1.1 below). Then if a string A over the alphabet had the numerical "codes" $a_0, a_1, \ldots, a_{n-1}$ for its symbols from left to right, then he assigned (essentially[24]) the code $\langle a_0, a_1, \ldots, a_{n-1}\rangle$ to A. In the modern literature this code of A is denoted by $\ulcorner A \urcorner$.

Further, he coded sequences of strings $A_0, A_1, \ldots, A_{n-1}$ —such as *proofs* within PA— by $\langle \ulcorner A_0 \urcorner, \ulcorner A_1 \urcorner, \ldots, \ulcorner A_{n-1} \urcorner\rangle$.

DEFINITION 7.1.1 (Alphabet and Axioms of PA). The *logical* part of the alphabet —the part is necessary to do just logic— is the standard one for any first-order classical logic *with equality* augmented with the *nonlogical* (or *mathematical*) symbols as follows: *one* constant, "0", *one* unary function "S", *two* binary functions "$+$" and "$\times$" and *one* binary predicate, "$<$".

The logical axioms are the A1 – A4 of 4.1.1 with A2 modified to allow any substitutable term t into x: $(\forall x)A \to A[t]$. The axioms of equality are two schemata, (1) $x = x$ for any choice of variable, and (2) $t = s \to (A[t] \leftrightarrow A[s])$ for any choice of terms t, s and formula A.

P1 $\neg Sx = 0$ for any variable x, where S is the successor function

P2 $Sx = Sy \to x = y$ for any variables x and y

P3 $x + 0 = x$ for any variable x

P4 $x + Sy = S(x + y)$ for any variables x and y

P5 $x \times 0 = 0$ for any variable x

P6 $x \times Sy = (x \times y) + x$ for any variables x and y

P7 $\neg x < 0$ for any variable x

[24] Actually Gödel used a different coding of number sequences; he used his "β-function" —cf. [14, 34, 38]— but the term "Gödel numbering" has since been attached to any one-to-one recursive coding scheme of finite number sequences.

P8 $x < Sy \leftrightarrow x < y \vee x = y$ for any variables x and y

P9 $x < y \vee x = y \vee y < x$ for any variables x and y

P10 (Induction Schema), for any formula $A[x]$ (cf. notation in Chapter 3):

$A[0] \wedge (\forall x)(A[x] \rightarrow A[Sx]) \rightarrow A$

PA is a *first-order* theory —or "applied *first-order* logic"— and as such satisfies first-order soundness and completeness.

There is another, *syntactic*, version of completeness —called *simple completeness*— for a theory such as PA, stating that for **every** *sentence* A of the theory the axioms can *prove* one of A or $\neg A$ (if the theory is consistent, then *exactly one* of these two), in symbols $\vdash_{PA} A$ or $\vdash_{PA} \neg A$.

In his *First Incompleteness Theorem* Gödel [14] proved that PA is NOT simply complete, and thus there is a sentence A —in fact, infinitely many— such that $\nvdash_{PA} A$ *and* $\nvdash_{PA} \neg A$.[25]

Such sentences are called *undecidable in the theory*.[26]

Enter the concept of the "*intended*" or "*standard*" model: In the case of PA it is the one that has as its domain of individuals the set $\mathbb{N}$, where the alphabet symbol "0" is interpreted as *the* $0 \in \mathbb{N}$, $<$ is interpreted as the "less than" order on $\mathbb{N}$, while "S" is interpreted as the successor function on the natural numbers and the binary "$+$" and "$\times$" are interpreted as "plus" and "times" on the natural numbers.

Now, the set of all axioms, the logical (trivially) and the mathematical P1–P10 are true in the standard model, a state of affairs Smullyan [36]) called the "correctness of PA".

By first-order *soundness*, all PA theorems are true in the standard model, thus a *semantic* formulation of the incompleteness theorem is that *there is a sentence S over the alphabet of PA, which is true in the standard model, but is* not *a PA theorem*! (for we cannot have both of the undecidable sentences A and $\neg A$ above being false).

[25] Technically, a key *assumption* for Gödel's theorem, clearly stated in [14], is that PA is consistent. Here we make two observations: 1) by definition, inconsistency is tantamount to "all formulae of the PA language are theorems", and thus would trivially be simply complete but totally useless as a theory; *and* 2) it is known that PA *is* consistent as the presence of a model —the standard one— establishes metamathematically.

[26] Not to be confused with undecidable —that is, not recursive— *sets* or *relations*.

In this semantic guise of the first incompleteness theorem a proof is almost easy and is grounded on just two fairly easily proved theorems of computability:

1. If Γ is a recursively axiomatized theory, then the set of Gödel numbers of closed theorems $\{\ulcorner A \urcorner : A \text{ is a closed theorem of } \Gamma\}$ is *semi-recursive* (also known as "recursively enumerable").

2. The *set of all sentences over the PA language* that are true in the standard model —the standard symbol for this set being CA ("complete arithmetic")— is NOT semi-recursive.

Thus, taking $\Gamma = PA$ and denoting the set of its closed theorems by $\mathcal{T}$, then, by 1 and 2, the latter set cannot equal CA. So, $\mathcal{T} \subsetneqq \text{CA}$ since by correctness of PA, trivially, $\mathcal{T} \subseteq \text{CA}$.

7.2. Arithmetical Completeness

The main tool in this section is Solovay's work [37]. (See AppendixC for a detailed account of Solovay's Theorem.) We build on [32] but also use two tools from [45], namely, a definition and a lemma in loc. cit., which appear modified below as 7.2.8 and 7.2.9 respectively. Our induction in the proof of 7.2.10 proceeds in its details differently.

Central in this section is the concept of the *realisation* of a modal logic in PA. Solovay defined the *realisation* or *interpretation* of the propositional logic GL in PA in [37]. In the case of GL this realisation, denoted by *, is a mapping from GL to PA that maps GL formulae to PA formulae. We write A^* for the image of A. The logic GL is the propositional sublogic of ML^3 obtained by removing (a) the axioms A1–A4 from 4.1.1, and (b) removing all first-order elements from the alphabet. Thus the only rule of inference in GL is MP, while necessitation is an admissible weak rule.

The mapping * is defined by induction on the formation (or complexity) of formulae of GL:

1. For $\bot$ in the language of GL we set $\bot^*$ to be the $\bot$ in the language of of PA.

2. For each propositional variable p, p^* is some sentence of PA that we pick as we please.

3. We let * commute with the classical propositional connectives. (E.g, $(A \wedge B)^*$ is $A^* \wedge B^*$.)

4. For the modal connective $\Box$ we let $(\Box A)^*$ be $\Theta_{PA}(\ulcorner A^* \urcorner)$, where $\Theta_{PA}(x)$ —Gödel's *provability predicate*— is defined to be $(\exists y) Pr_{PA}(y, x)$. As for $Pr_{PA}(y, x)$, this is the recursive relation "the formula of Gödel number x has a proof of Gödel number y".

Solovay proved both *arithmetical soundness* and *arithmetical completeness* of GL under *:

- Soundness: If $\vdash_{GL} A$ then $\vdash_{PA} A^*$, for any *.
- Completeness: If, *for all* realisations *, we have $\vdash_{PA} A^*$, then it is also the case that $\vdash_{GL} A$.

Note that as p^* is arbitrarily defined (item 2. above) we have infinitely many realisation mappings.

The central result in this thesis is the proof of the arithmetical completeness of the predicate modal logic ML^3.

We derive the standard concept of the realisation mapping, from the language of ML^3 to the language of PA, as it applies to predicate logics where $\Box A$ is always a sentence. The definition is via the clauses below that are those of the list 1.–4. above with some adaptation to the predicate case (cases 2. and 3.), plus two new definition clauses, 5. and 6.:

DEFINITION 7.2.1 (Definition of a realisation of ML^3 in PA).

1. For $\bot$ in the language of ML^3 we set $\bot^*$ to be the $\bot$ in the language of PA.
2. For each atomic (other than $\bot$) A, A^* is some sentence of PA that we pick as we please.
3. We let * commute with the classical propositional connectives *and with the universal quantifier.*
4. For the modal connective $\Box$ we let $(\Box A)^*$ be $\Theta_{PA}(\ulcorner A^* \urcorner)$.
5. A^* over the language of PA has the same free variables as A over the language of ML^3.
6. * commutes with the operation of substitution of variables for variables, that is, if $A^*[x]$ is $B[x]$ in the language of PA, then $A^*[y]$ is $B[y]$.

Soundness and completeness with respect to this realisation is as in the two bullets above, replacing "GL" by "ML^3".

At the end of this introduction to our main topic, we state the main theorem of this thesis below, and embark on its proof.

THEOREM 7.2.2 (Main Theorem). *ML^3 is arithmetically complete in some recursive extension $\mathcal{T}$ of PA in the sense that, for any closed A over the language of ML^3, if all arithmetical realisations A^* of A are provable in $\mathcal{T}$, then A is provable in ML^3.*

As was done in [37] (for GL), we prove 7.2.2 contrapositively: Thus, assume $\mathrm{ML}^3 \nvdash A$, for some fixed modal *sentence* A over the language L of ML^3, and find an arithmetical realisation in $\mathcal{T}$ such that $\nvdash_{\mathcal{T}} A^*$.

The first phase of this plan is to build a finite, irreflexive and transitive Kripke model $\mathcal{M} = \left(\left\langle W^f, \widehat{R}, \alpha_0 \right\rangle, \{(\mathbb{N}, \Vdash_S) : \alpha_i \in W^f\}\right)$ for $\mathrm{ML}^3 \cup \{\neg A\}$, therefore one where

$$\alpha_0 \nVdash_S A \tag{$\dagger$}$$

This was done in 6.3.8 above.

The second phase is to apply Solovay's technique [37] to embed $\mathcal{M}$ in an appropriate $\mathcal{T}$ —*which is a* finite *extension of PA that we define below*— and propose *a specific* arithmetical *realisation* *, for which we will prove that $\mathcal{T} \nvdash A^*$. This will complete the proof of 7.2.2.

An *a priori* requirement of the embedding is that the worlds α_i (cf. 6.3.5) make sense in the language of PA, thus we rename them into numbers.

$$W^f = \{1, 2, 3, \ldots, n\}$$

where "$i + 1$" stands for "α_i".

For technical reasons Solovay adds a new world named 0 —in our case with $M_0 = \mathbb{N}$— and modifies $\mathcal{M}$ to $\mathcal{M}^0$, by modifying:

- $\widehat{R}$ into $\widehat{R}^0 = \widehat{R} \cup \{(0, i) : 1 \leq i \leq n\}$
- the forcing relation $\Vdash_S$ into $\Vdash_{S^0}$ by letting $0 \Vdash_{S^0} X$ iff $1 \Vdash_S X$, while $i \Vdash_{S^0} X$ iff $i \Vdash_S X$, for $1 \leq i \leq n$.

- $W^{f,0} = \{0, 1, 2, 3, \ldots, n\}$

It is this $\mathcal{M}^0$ that Solovay embeds into PA (or extension $\mathcal{T}$). Below we list the Solovay's lemmata that, interestingly, *can be used here as is* without reference to their complex proofs (not so in [2, 23]) and without any reference to symbols and concepts used in their proofs. For simplicity of *use* and exposition —and to ascertain independence from the inner workings of Solovay's proofs— many authors ([2,4,45]) use the abbreviations S_k or σ_k for the formal sentence (in PA) "$\mathbf{l} = \widetilde{k}$" that is pervasive in [37], where $\widetilde{k}$ is the formal counterpart in PA —a *numeral*— of the number $k \in \mathbb{N}$ and $\mathbf{l}$ denotes a formal term that is the *limit* of Solovay's "function h" whose outputs are in $W^{f,0}$, a function whose definition or properties we need to know nothing towards applying the lemmata below to prove 7.2.2.

LEMMA 7.2.3 (Solovay's Lemmata). *Let $\mathcal{T}$ be some consistent recursive extension of PA over a finite extension of the PA language. There are sentences S_i, for $0 \leq i \leq n$, of the language, such that*

(1) *For all $i \neq j$, $\vdash_{\mathcal{T}} \neg S_i \vee \neg S_j$.*

(2) *For $0 \leq i \leq n$, $\mathcal{T} + S_i$ is consistent.*

(3) *If $i \widehat{R}^0 j$, then $\vdash_{\mathcal{T}} S_i \rightarrow \neg\Theta_{\mathcal{T}}(\ulcorner \neg S_j \urcorner)$, where $\Theta_{\mathcal{T}}$ is the provability predicate for $\mathcal{T}$.*

Under the given assumptions, [37] formulated this item as the equivalent $\vdash_{\mathcal{T}} S_i \rightarrow Cons_{\mathcal{T}+S_j}$. In words, $\mathcal{T}$ proves the formalised in $\mathcal{T}$ consistency of $\mathcal{T} + S_j$ from premise S_i.

(4) *If $i > 0$, then $\vdash_{\mathcal{T}} S_i \rightarrow \Theta_{\mathcal{T}}(\ulcorner \bigvee_{i \widehat{R}^0 j} S_j \urcorner)$.*

As in [45] we will work with a specific finite consistent extension $\mathcal{T}$ of PA rather than PA itself. Towards obtaining this theory, we build consistent sets of *classical* formulae $\mathcal{C}_i$ (7.2.5 below) as follows.

We note that while $i \cap S(A)$ is consistent it is *not a maximal* consistent finite subset of $S(A)$ since i contains only sentences. Thus if $X(y)$ is in $S(A)$ —as a result of the presence of $(\forall y)X$ as a closed subformula of A— it is not in i ($= \alpha_{i-1}$). On the other hand, if $(\forall y)X$ is consistent with ML^3,[27] then so is $X(y)$ and vice versa by virtue of $\vdash (\forall y)X \rightarrow X(y)$ absolutely (axiom A2) and $X(y) \vdash (\forall y)X$. Thus we define

[27] Meaning that $(\forall y)X \nvdash_{ML^3} \perp$, where $\perp$ is a convenient formula that represents a contradiction.

DEFINITION 7.2.4. For each $1 \leq i \leq n$, $S^i_{\max}(A)$ denotes a *maximal consistent subset of* $S(A)$ that contains $i \cap S(A)$ $(= \alpha_{i-1} \cap S(A))$.[28] □

Such an $S^i_{\max}(A)$ along with a $\forall X$ that it might contain will also contain all formulae obtained from $\forall X$ by stripping one $(\forall u)$ at a time, from left to right, from the prefix $\forall$ of X (axiom A2).

DEFINITION 7.2.5. We next define a set of classical formulae $\mathcal{C}_i$, for each $1 \leq i \leq n$.

(1) If $X \in S^i_{\max}(A)$ is a classical first-order formula, then X is *transformed* into itself (no change), and is added to $\mathcal{C}_i$ under the name $X^{t,i}$.

(2) If $X \in S^i_{\max}(A)$ contains at least one $\Box$, then every *top level occurrence* of $\Box B$ *in* X is changed to $\top$ iff $\Box B \in i$, else it is changed to $\bot$.[29] The transformed formula X —again given the name $X^{t,i}$— is placed in $\mathcal{C}_i$. □

Remark 7.2.6. "t" is for "transformed" formula. But why the extra superscript i? Because the same X may appear in i and j, for $i \neq j$. But some top level subformula $\Box B$ of X may be in i but not in j. This results in having two distinct transforms $X^{t,i}$ and $X^{t,j}$. □

PROPOSITION 7.2.7. $\mathcal{C}_i$ is consistent iff $S^i_{\max}(A)$ is consistent.

PROOF: Let $X \in S^i_{\max}(A)$. Note that, if the subformula $\Box B$ of X is in i, then $i \vdash \Box B \equiv \top$[30] while if $\Box B \notin i$, then $\neg\Box B$ is in i by maximality, hence $i \vdash \Box B \equiv \bot$.[31]

Now let $\Box B \in S^i_{\max}(A)$. Then the first $\vdash$-statement above is refined to $S^i_{\max}(A) \vdash \Box B \equiv \top$. In the opposite case $\neg\Box B$ is in i (and in $S(A)$ by Definitition 6.3.4) and thus in $S^i_{\max}(A)$. Hence $S^i_{\max}(A) \vdash \Box B \equiv \bot$.

Therefore $S^i_{\max}(A) \vdash X \leftrightarrow X^{t,i}$ since $X^{t,i}$ is obtained by a finite sequence of replacing "equivalents by equivalents" according to the preceding paragraph. Thus, $\mathcal{C}_i$ proves $\bot$ iff $S^i_{\max}(A)$ proves $\bot$. □

[28] Such maximal consistent subsets trivially exist by finiteness of $S(A)$.

[29] Case of $\neg\Box B$ being in i. Incidentally, if X contains the subformula $\Box(\ldots\Box C\ldots)$ at the top level it is clear that there is no point to replace $\Box C$ by $\top$ or $\bot$.

[30] $i \vdash \Box B$ and tautological implication.

[31] Since $i \vdash \neg\Box B$.

Now, each $S^i_{\max}(A)$ *is* consistent, hence *each* $\mathcal{C}_i$ is also a consistent finite set of (*classical*) formulae over the language $L(\mathbb{N})$.

Note that the formulae X of the classical sets $\mathcal{C}_i$ with parameters in $\mathbb{N}$ can each be realised in the language of PA (cf. also [21, Vol. II] and [20, 25]) as a true formula in the standard model. Indeed, add all the *finitely* many predicate letters found in $\mathcal{C}_i$ to the language of PA and also replace each parameter $\mathbf{k}$ (imported constant, 6.2.1) that occurs in every such X into the numeral $\widetilde{k}$ to obtain a formula $re_i(X)$ in the language of PA. *Both* $\mathbf{k}$ *and* $\widetilde{k}$ *are interpreted as* $k \in \mathbb{N}$. We denote by $re_i(\mathcal{C}_i)$ the set $\{re_i(X) : X \in \mathcal{C}_i\}$.

It follows that each set $re_i(\mathcal{C}_i)$ is *consistent with* PA since the latter's standard model is also a model of $re_i(\mathcal{C}_i)$ and thus of PA + $re_i(\mathcal{C}_i)$ as well.

Therefore, for each $i = 1, \ldots, n$, we can consistently add to PA the new axiom

$$\mathscr{A}_i \overset{Def}{\leftrightarrow} \left(\bigwedge_{X \in re_i(\mathcal{C}_i)} X \right)$$

We define

$$\mathcal{T} \overset{Def}{=} PA + \{\mathscr{A}_1, \ldots, \mathscr{A}_n\}$$

Referring to Definition 7.2.1 we now state:

DEFINITION 7.2.8 (Arithmetical realisation; *initialisation*).

Let B be any *atomic* subformula of A (other than $\perp$), where A was fixed at the outset of this section (cf. (†) p.42). Being atomic it is classical.

Then for the *basis* of the realisation * (cf. 7.2.1) we set ([45]),[32]

$$B^* \overset{Def}{\leftrightarrow} \bigvee_{\substack{1 \le j \le n \\ j \Vdash \forall B}} S_j \wedge re_j(B^{t,j}) \tag{1}$$

If the $\bigvee$ is empty, then we set B^* to be a simple expression equivalent to $\perp$, say, $\neg \bigwedge_{1 \le i \le m} u_i = u_i$, where $u_1, u_2, \ldots, u_m$ are all the free variables of B and thus of B^*. Of course, unlike ML^3 in this thesis, $\mathcal{T}$ is a logic *with* equality. □

The following useful lemma is stated in Yavorsky [45] without proof. Our proof is the following.

[32] Recall the renaming of α_j as $j + 1$, on p.42.

LEMMA 7.2.9. $\vdash_{\mathcal{T}} S_i \to (B^* \leftrightarrow re_i(B^{t,i}))$ *for any* classical *first-order subformula* B *of* A, *and* $1 \leq i \leq n$.

PROOF: We do induction on the *classical* complexity of B (number of $\neg, \wedge$ and $\forall$ connectives).

First, since S_i is a sentence, invoking the deduction theorem

$$\text{we need to prove instead } \vdash_{\mathcal{T}+S_i} B^* \leftrightarrow re_i(B^{t,i}) \tag{2}$$

We now proceed with our induction on classical formulae B:

1. If B is $\bot$, then so is $re_i(B^{t,i})$. Thus, since $\bot^*$ is $\bot$, (2) is trivial.

2. B *is atomic* (*Basis*): Having S_i as a hypothesis in (2), tautological implication yields from (1),

$$\vdash_{\mathcal{T}+S_i} B^* \leftrightarrow re_i(B^{t,i}) \vee \bigvee_{\substack{j \neq i \\ j \Vdash \forall B}} S_j \wedge re_j(B^{t,j})$$

   Note that by 7.2.3(1), we have $\vdash_{\mathcal{T}+S_i} \neg S_j$ for $j \neq i$. Thus by tautological implication the "$\bigvee$" part above drops out (is provably equivalent to $\bot$). We have proved the Basis step.

   We omit the cases of Boolean connectives as trivial but sample the equally trivial case of the $\forall$ connective below.

3. B is $(\forall x)D$. By I.H. $\vdash_{\mathcal{T}+S_i} D[x]^* \leftrightarrow re_i(D^{t,i}[x])$. By the equivalence theorem, $\vdash_{\mathcal{T}+S_i} (\forall x)D^* \leftrightarrow (\forall x)re_i(D^{t,i})$. But $((\forall x)D)^*$ is $(\forall x)D^*$ by the definition of * (7.2.1, (3)) while, by the definition of re_i, $re_i\big((\forall x)D^{t,i}\big)$ is $(\forall x)re_i(D^{t,i})$.

□

The proof of the Main Lemma below will use Löb's "derivability conditions" (DC) 1 and 2 which we list below for the record (cf. [38] for their rather lengthy proofs).

DC 1 If $\vdash_{\mathcal{T}} A$, then $\vdash_{\mathcal{T}} \Theta(\ulcorner A \urcorner)$.

DC 2 $\vdash_{\mathcal{T}} \Theta(\ulcorner A \to B \urcorner) \to \Theta(\ulcorner A \urcorner) \to \Theta(\ulcorner B \urcorner)$.

LEMMA 7.2.10 (Main Lemma). *Having got a finite Kripke model,* $\mathcal{M}^0$, *of p.43), of* $n+1$*-nodes such that* $1 \nVdash_{S^0} A$ *(cf.* (†) *on 42), where "1" is* α_0 *and "n" is* α_{n-1} *and* A *is closed, we will prove, for* every closed *subformula* X *of* A, *and for all* $1 \leq i \leq n$, *that*

(1) *If* $i \Vdash_{S^0} X$, *then* $\vdash_{\mathcal{T}} S_i \to X^*$.

(2) *If* $i \nVdash_{S^0} X$, *then* $\vdash_{\mathcal{T}} S_i \to \neg X^*$.

PROOF: Induction on the complexity of the modal *sentence* X. Throughout, by the deduction theorem, we routinely replace the tasks "$\vdash_{\mathcal{T}} S_i \to \ldots$" by the tasks "$\vdash_{\mathcal{T}+S_i} \ldots$"

1. X is atomic.

   (a) Verify (1) of the lemma. So we have $i \Vdash_{S^0} X$. Hence (by 7.2.9) $\vdash_{\mathcal{T}+S_i} X^* \leftrightarrow re_i(X^{t,i})$. But $re_i(X^{t,i})$ is a conjunct of an axiom of $\mathcal{T}$ thus $\vdash_{\mathcal{T}} re_i(X^{t,i})$. By tautological implication, $\vdash_{\mathcal{T}+S_i} X^*$.

   (b) Verify (2) of the lemma. So $i \nVdash_{S^0} X$, thus in (1) of 7.2.8, the disjunct $S_i \wedge re_{i-1}(X^{t,i-1})$ is missing. By item 2. in the proof of 7.2.9 we have $\vdash_{\mathcal{T}+S_i} X^* \leftrightarrow \bot$, that is, $\vdash_{\mathcal{T}+S_i} \neg X^*$.

   The interesting induction steps are for X of the form $\Box B$ or $(\forall x)B$.

2. X is $\Box B$.

   (1) *of the Lemma.* Assume $i \Vdash_{S^0} \Box B$. Then for all j such that $i\widehat{R}^0 j$ it is $j \Vdash_{S^0} \forall B$. By I.H.[33] and definition by cases,

$$\vdash_{\mathcal{T}} \bigvee_{i\widehat{R}^0 j} S_j \to (\forall B)^*$$

   Applying DC1 then DC2 followed by modus ponens,

$$\vdash_{\mathcal{T}} \Theta(\ulcorner \bigvee_{i\widehat{R}^0 j} S_j \urcorner) \to \Theta(\ulcorner (\forall B)^* \urcorner) \qquad (*)$$

   By 7.2.3(4) $\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner \bigvee_{i\widehat{R}^0 j} S_j \urcorner)$ and hence, by $(*)$,

$$\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner (\forall B)^* \urcorner) \qquad (**)$$

   Now $\vdash \forall B \to B$ (absolutely) and also $\vdash_{\mathcal{T}} (\forall B)^* \to B^*$ since $(\forall B)^*$ is $\forall(B^*)$. Hence, by DC1 and DC2, $\vdash_{\mathcal{T}} \Theta(\ulcorner (\forall B)^* \urcorner) \to \Theta(\ulcorner B^* \urcorner)$.

[33] We remind the reader that as in [32] $\Box B$ is more complex than $\forall B$.

This and tautological implication from $(**)$ yields

$$\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner B^* \urcorner)$$

Noting that $(\Box B)^*$ is $\Theta(\ulcorner B^* \urcorner)$, this case is done.

(2) *of the Lemma.* Assume $i \nVdash_{S^0} \Box B$. Then for *some* j such that $i\widehat{R}^0 j$ it is $j \nVdash_{S^0} \forall B$. *We pick one such* j.

By I.H.

$$\vdash_{\mathcal{T}} S_j \to \neg(\forall B)^*$$

hence $\vdash_{\mathcal{T}} (\forall B)^* \to \neg S_j$. By DC1 and DC2, $\vdash_{\mathcal{T}} \Theta(\ulcorner (\forall B)^* \urcorner) \to \Theta(\ulcorner \neg S_j \urcorner)$, hence

$$\vdash_{\mathcal{T}} \neg\Theta(\ulcorner \neg S_j \urcorner) \to \neg\Theta(\ulcorner (\forall B)^* \urcorner) \qquad (\ddagger)$$

By 7.2.3(3), $i\widehat{R}^0 j$ yields $\vdash_{\mathcal{T}} S_i \to \neg\Theta(\ulcorner \neg S_j \urcorner)$. Therefore, a tautological implication using this and $(\ddagger)$ derives

$$\vdash_{\mathcal{T}} S_i \to \neg\Theta(\ulcorner (\forall B)^* \urcorner) \qquad (***)$$

By successive applications of axiom A7 of ML^3 we obtain $\vdash_{ML^3} \Box B \to \Box\forall B$, hence (by definition of * and arithmetical soundness, the latter not proved in this thesis, but see [32]), $\vdash_{\mathcal{T}} (\Box B)^* \to (\Box\forall B)^*$, that is, $\vdash_{\mathcal{T}} \Theta(\ulcorner B^* \urcorner) \to \Theta(\ulcorner (\forall B)^* \urcorner)$. From $(***)$ and the preceding we now get $\vdash_{\mathcal{T}} S_i \to \neg\Theta(\ulcorner B^* \urcorner)$ by tautological implication, that is, $\vdash_{\mathcal{T}} S_i \to \neg(\Box B)^*$.

3. X is $(\forall x)B$. If the quantification is not redundant, then the subformula B is not a sentence and the I.H. does not apply to it. Thus we proceed using 7.2.9 instead.

(I) $(\forall x)B$ is classical. Thus

$$\vdash_{\mathcal{T}+S_i} (\forall x)B^* \leftrightarrow re_i((\forall x)B) \qquad (0)$$

(a) Now, if $i \Vdash_{S^0} (\forall x)B$, then $\vdash_{\mathcal{T}} re_i((\forall x)B)$. Tautological implication and (0) yield $\vdash_{\mathcal{T}+S_i} (\forall x)B^*$.

(b) If $i \nVdash_{S^0} (\forall x)B$, then $(\forall x)B$ is false in the world i, hence the true $\neg(\forall x)B$ is in $S^i_{\max}(A)$. Thus $re_i\big(\neg(\forall x)B\big)$ is a conjunct of an axiom of $\mathcal{T}$ and therefore $\vdash_{\mathcal{T}} re_i(\neg(\forall x)B)$, i.e., $\vdash_{\mathcal{T}} \neg re_i((\forall x)B)$.

(0) now yields $\vdash_{\mathcal{T}+S_i} \neg(\forall x)B^*$.

(II) $(\forall x)B$ is *not* classical.

(a) Assume $i \Vdash_{S^0} (\forall x)B$.

- Let $\Box C$ be a topmost occurrence in $(\forall x)B$ and $\Box C \in S^i_{\max}(A)$.
  Let B' be B with said occurrence of $\Box C$ replaced by $\top$. Since $i \Vdash_{S^0} (\forall x)B$ iff $i \Vdash_{S^0} (\forall x)B'$ the I.H. yields
  $$\vdash_{\mathcal{T}+S_i} \big((\forall x)B'\big)^* \tag{1}$$
  The I.H. also yields $\vdash_{\mathcal{T}+S_i} \big(\Box C\big)^*$, hence $\vdash_{\mathcal{T}+S_i} \big(\Box C\big)^* \leftrightarrow \top$ (recall that $\top^*$ is by definition $\top$). From the latter and the equivalence theorem we get $\vdash_{\mathcal{T}+S_i} (\forall x)B^* \leftrightarrow \big((\forall x)B'\big)^*$ and we are done by (1).
- Let $\Box C$ be a topmost occurrence in $(\forall x)B$ and $(\neg\Box C) \in S^i_{\max}(A)$. This is entirely analogous with the above, but note that we replace here $\Box C$ by $\bot$ on the ML^3 side and by $\bot^*$ on the $\mathcal{T}$ side.

(b) Assume $i \nVdash_{S^0} (\forall x)B$.

- Let $\Box C$ be a topmost occurrence in $(\forall x)B$ and $\Box C \in S^i_{\max}(A)$.
  Let B' be B with said occurrence of $\Box C$ replaced by $\top$. Since $i \nVdash_{S^0} (\forall x)B$ iff $i \nVdash_{S^0} (\forall x)B'$ the I.H. yields
  $$\vdash_{\mathcal{T}+S_i} \neg\big((\forall x)B'\big)^* \tag{2}$$
  The concluding paragraph of this subcase proceeds exactly as in bullet one of (I): we have $\vdash_{\mathcal{T}+S_i} (\forall x)B^* \leftrightarrow \big((\forall x)B'\big)^*$ but this time it is (2) that yields $\vdash_{\mathcal{T}+S_i} \neg(\forall x)B^*$.
- The subcase where a topmost occurrence of $\Box C$ in $(\forall x)B$ satisfies $(\neg\Box C) \in S^i_{\max}(A)$ does not offer any new insights.

□

Proof of the main theorem. By 7.2.10, since A is a subformula of itself and $1 \nVdash_{S^0} A$ we have $\vdash_{\mathcal{T}} S_1 \to \neg A^*$. By Lemma 7.2.3(2) $\mathcal{T} + S_1$ is consistent, hence so is $\mathcal{T} + \neg A^*$.[34] Thus $\nvdash_{\mathcal{T}} A^*$. □

[34] If $\vdash_{\mathcal{T}+\neg A^*} \bot$, then $\vdash_{\mathcal{T}} A^*$ thus also $\vdash_{\mathcal{T}+S_1} A^*$ contradicting the consistency of $\mathcal{T} + S_1$.

Chapter 8

Concluding Note

This thesis adds to the example of QGL^b ([45]) another predicate modal logic, ML^3, that is arithmetically complete. It also contains an Appendix with a more complete and mathematically rigorous proof of Solovay's original Lemmata than any that we are aware of in the literature, including Solovay's own paper [4, 10, 23, 35, 37]. Unlike QGL^b the ML^3 does not have necessitation as a primary rule and as a result has the added desirable attribute that some of its metatheory can be investigated without Gentzenisation, using formulators — [33, 41]. The second of the preceding two references contains a result of independence for axiom A7, which implies that *its presence is essential* —in fact, removing it we *get an arithmetically incomplete logic*. The origins of QGL^b and ML^3 are quite distinct, as the former was aimed to answer "are there arithmetically complete first-order modal logics?" while the origin of ML^3 (via its predecessor M^3) was to build a modal first-order logic that can effectively simulate classical first-order equational proofs. Thus the former chose the "*opaque*" $\Box$ to avoid known negative results —that hinge on the presence of a "transparent" $\Box$— towards arithmetical completeness, while the latter chose this very same feature for a *totally different design reason*: to enable M^3 and ML^3 to simulate using $\Box$ the classical $\vdash$ of a logic where $A \vdash (\forall x)A$ is an unconstrained rule. This is carefully explained in [42, 43] where also the choice of A7 is explained as the modal counterpart of the classical $A \vdash (\forall x)A$.

A7 appears to have been adopted without any discussion in [45]. In particular loc. cit. seems to be unaware of this axiom's indispensability in the context of the remaining axioms, as such a fact was proved

in [41].

There is a question whether the (finite) extension of PA that we used can be either 1) compressed to PA without affecting the proof of arithmetical completeness of ML^3, or 2) whether a direct proof that ML^3 is arithmetically complete with realisations being *in PA*. In other words, is ML^3 arithmetically complete in the standard PA? The answer is not at all obvious. This is an open problem worth investigating. Artemov and Dzhaparidze proposed an arithmetically complete predicate modal logic [2], where the realisations happens in PA *and* the $\Box$ is transparent. But they assumed unrealistic Kripke models that have *finite domains* such that if the world w accesses world w' then the domain of w is a subset of the domain of w'.

Appendix A

Proof of the Main Semantic Lemma (6.2.2)

We first introduce some tools.

THEOREM A.0.1 (The Variant Theorem). *Let z be fresh for $(\forall x)A$. Then ML^3 proves the equivalence $(\forall x)A[x] \equiv (\forall z)A[z]$.*

PROOF: The proof of this elementary and standard result can be found in all of [34, 38, 39]. □

THEOREM A.0.2 (The Metatheorem on New Constants, [34, 38]). *Let c be a new constant that* does not occur in the axioms of a theory $\mathcal{T}$, *which includes the logical axioms C.2.1 of ML^3.*

Let us add c to the theory $\mathcal{T}$ without affecting its nonlogical (mathematical) axioms.

Then, for any formula $A[x]$, if $\mathcal{T} \vdash A[c]$ and z is a variable that does not occur in $A[c]$ as either free or bound, we will also have that $\mathcal{T} \vdash A[z]$.

PROOF: We do induction on the generation of theorems $A[c]$:

- (*Basis*) $A[c]$ is an axiom. Then so is $A[z]$ as determined by the *forms* of $A[c]$ and $A[z]$ —they are the same. The newness of z enters in the argument regarding the form in the cases of A2, A3, A4 of 4.1.1 where capture of z (after it replaces c) might have distorted the form or caused non substitutability. *Such capture is impossible.* Of course, the nonlogical axioms of $\mathcal{T}$ are also immune: c does not occur in them.

- $A[c]$ was derived by MP from $X[c]$ and $X[c] \to A[c]$. By the I.H. $X[z]$ and $X[z] \to A[z]$ are $\mathcal{T}$-theorems thus so is $A[z]$ by MP.
- $A[c]$ is $(\forall w)B[w, c]$, and $\mathcal{T} \vdash B[w, c]$. By the I.H. $\mathcal{T} \vdash B[w, z]$ and thus $\mathcal{T} \vdash (\forall w)B[w, z]$ by generalisation.

□

Suppose that we have $\vdash_\Gamma (\exists x)A[x]$ and we want to prove B. Informally we say "**let** c be a 'constant' not in B or in our hypotheses, which makes $A[c]$ 'true' —in effect, adding the latter formula as a *new* hypothesis".

Then we embark on proving B using $A[c]$ along the other hypotheses.

Once we are done, we conclude "$\vdash_\Gamma (\exists x)B$, that is, $A[c]$ has dropped out from the hypotheses list".

Why does this help? Because in practice $A[c]$ may have some Boolean structure that we can hopefully take advantage of. The "packed" $(\exists x)A[x]$ is hard to use, much as $(\forall x)A[x]$ which we quickly strip of its $\forall$ using Ax2 (4.1.1) and MP.

The following formalises the $\exists$-elimination process.

Warning. Note the "**let**" above. We do not *have* this c! In $\exists$-elimination $A[c]$ is **a new hypothesis**, unlike the case of going from $(\forall x)A[x]$ to $A[t]$.

COROLLARY A.0.3 (Elimination of $\exists$). Suppose that we have $\vdash_\Gamma (\exists x)A[x]$ and we want to prove B.

If we pick a new constant c that is not in any of Γ or $(\exists x)A[x]$ or B and show

$$\vdash_{\Gamma + A[c]} B \tag{1}$$

then we can conclude that $\vdash_\Gamma B$.

PROOF: By the deduction theorem (1) implies $\vdash_\Gamma A[c] \to B$.

Let z be fresh (not in Γ or $(\exists x)A[x]$ or B). By A.0.2 we have also $\vdash_\Gamma A[z] \to B$ where z is a variable that does not occur in our hypotheses or B. By $\exists$-monotonicity[35] we obtain $\vdash_\Gamma (\exists z)A[z] \to (\exists z)B$ hence

$$\vdash_\Gamma (\exists z)A[z] \to B \tag{2}$$

[35]This is the $A \to B \vdash (\exists)A \to (\exists x)B$ of pure logic (cf. [38, 39]).

since B contains no z. Now the bound variable variant theorem of pure logic yields (on the freshness assumption for z) $\vdash (\exists z)A[z] \leftrightarrow (\exists x)A[x]$ hence the hypothesis $\vdash_\Gamma (\exists x)A[x]$ yields $\vdash_\Gamma (\exists z)A[z]$ by tautological implication.

MP and (2) conclude our proof. □

PROOF: (of 6.2.2)

1. By initialisation $\Gamma_0 = \mathcal{T}$.

2. We prove first that Γ_n is consistent for all $n \geq 0$. For the Basis, Γ_0 is so by initialisation.

   Suppose now (I.H.) that Γ_n is consistent for some fixed n.

   The case for Γ_{n+1}:

   We obtain Γ_{n+1} via Δ_n.

   - If Δ_n is given by the top case on p.27, then Δ_n is consistent, because otherwise, by 4.3.4, we have $\Delta_n = \Gamma_n \cup \{A_n\} \vdash \bot$, hence (4.3.5) $\Gamma_n \vdash \neg A_n$, contradicting the condition of the top case.
   - If Δ_n is given by the bottom case on p.27, then $\Delta_n = \Gamma_n$ is again consistent by the I.H. for Γ_n.

   $$\text{We established that } \Delta_n \text{ is consistent} \tag{1}$$

   Now, according to the construction on p.27, Γ_{n+1} is Δ_n (bottom case) —hence is consistent by (1)— or it is

   $$\Delta_n \cup \{A[\mathbf{m}]\} \text{ (top case)} \tag{2}$$

   where we have

   $$\Delta_n \vdash F_{n+1} \tag{3}$$

   F_{n+1} standing for $(\exists x)A[x]$ and $\mathbf{m}$ does not occur in any of A_0, $A_1, \ldots, A_n, F_0, F_1, \ldots, F_{n+1}$ —in other words, *this constant is not in any of the nonlogical axioms of* Δ_n.

   We conclude the proof of this item —number 2.— of 6.2.2 by arguing the consistency of the set in (2).

   If it is not, then (by 4.3.4 followed by 4.3.5) $\Delta_n \vdash \neg A[\mathbf{m}]$.

As $\mathbf{m}$ is fresh for Δ_n, picking a fresh variable z we have $\Delta_n \vdash \neg A[z]$, hence (by generalisation) $\Delta_n \vdash (\forall z)\neg A[z]$.

The latter says $\Delta_n \vdash \neg(\exists z)A[z]$ thus along with (3) —and A.0.1— we have that Δ_n is inconsistent, contradicting (1).

Thus we finished our induction that F_n is consistent for all n.

But now so is Γ for if not, then $\Gamma \vdash \bot$ and thus $\Gamma_n \vdash \bot$ for some n —because trivially, $\Gamma_n \subseteq \Gamma_{n+1}$ for all n, and a proof from Γ involves finitely many hypotheses; eventually all are in some Γ_n.

3. Let B be a sentence. Then B is an A_m and $\neg B$ is an A_k in the enumeration of all sentences (2. of p.27). Now, if neither B nor $\neg B$ were placed in Γ it must be that the top conditions failed in (cf. definition of the Γ-sequence on p.27)

$$\Delta_m = \begin{cases} \Gamma_m \cup \{A_m\} & \text{if } \Gamma_m \nvdash \neg A_m \\ \Gamma_m & \text{otherwise} \end{cases}$$

and

$$\Delta_k = \begin{cases} \Gamma_k \cup \{A_k\} & \text{if } \Gamma_k \nvdash \neg A_k \\ \Gamma_k & \text{otherwise} \end{cases}$$

Thus we have $\Gamma_m \vdash \neg B$ and $\Gamma_k \vdash B$. Since $\Gamma_m \cup \Gamma_k \subseteq \Gamma$, this renders Γ inconsistent. A contradiction to what we proved above.

4. (cf. [38]) Since the proof involves finitely many formulae from Γ, there is an $n \geq 0$ such that $\Gamma_{n+1} \supseteq \Delta_n \vdash (\exists x)A$. Since $(\exists x)A$ occurs in the enumeration of (3) (p.27) infinitely often, $(\exists x)A$ is the same as a F_{k+1} for some $k \geq n$. But then $\Delta_k \vdash (\exists x)A$ as well. Hence, $A[\mathbf{m}]$ is added to Γ_{k+1} as a Henkin axiom, and we are done.

□

Appendix B

The Recursion Theorem

Solovay's Lemmata (cf. 7.2.3 and C.5.2) are about the *limits* of a function "h" he defined in his paper ([37]). This h is defined on natural number inputs, and its outputs are among the nodes of the finite graph that represents the *finite*, *irreflexive* and *transitive* (hence also *reverse well-founded*) relation R of a finite Kripke frame (cf. 5.0.1).

Solovay relies on the recursion theorem of computability, ostensibly *proved within PA* (but he does not give a proof or reference), to construct h. Then he proceeds to prove very informally the properties of the limits of h albeit promising (without proof) that his arguments can be formalised within PA.

As his lemmata are central to his completeness result, as the reader who read the preceding parts of this thesis (especially the proof of 7.2.10) will have noted, and since moreover these lemmata are part of the (syntactic) *proof theory* of modal logics we felt obliged to give a closer approximation to *syntactic* formal proofs here, as much as possible that is consistent with a Masters level thesis.

Thus, following [4, 10] we will use Logic's "version" of the recursion theorem —know as the Gödel "*diagonalisation lemma*" ([14]) or, more fully attributed, the Gödel-Carnap ([5]) "diagonalisation lemma"— in defining h rather than relying on computability theory's *recursion theorem*. Nevertheless, we introduce both the latter and the diagonalisation lemma in this Appendix.

B.1. A brief review of elementary Computability

As a general and complete reference for this section we cite three sources [34, 38, 40].

Computability or Recursion Theory has at its centre the relation

$$\phi_e^{(n)}(\vec{x}_n) = y$$

in Rogers' ([30]) "ϕ-notation" or

$$\{e\}^{(n)}(\vec{x}_n) = y$$

in Kleene's ([25]) notation.

Either notation states that "a *program* that has *numerical code* $e \in \mathbb{N}$, when it is presented with input $\vec{x}_n$ from $\mathbb{N}^n$, it returns the value $y \in \mathbb{N}$".

The details of the programming languages where the program (of code) e is written are unimportant in this brief overview —popular such languages being the Turing Machine (TM) and the Unbounded Register Machine (URM)— as long as the two following *properties* of programs, *Smn* and *universal function*, below hold:

A good theory of computability allows the computation of non total functions, that is, *it is not always the case* that

$$(\forall \vec{x}_n)(\exists y)\phi_e^{(n)}(\vec{x}_n) = y$$

A function $f : \mathbb{N}^n \to \mathbb{N}$ is (partial) computable iff there is an e such that

$$(\forall \vec{x}_n, y)(f(\vec{x}_n) = y \leftrightarrow \phi_e^{(n)}(\vec{x}_n) = y) \tag{1}$$

DEFINITION B.1.1 (Partial Computable Functions). A function such as f in (1) is called *partial computable*. If it is defined for all $\vec{x}_n$, then it is called *computable*, or *total* computable if emphasis is called for.

Thus, f of n arguments is (partial) computable —by definition— iff, for some $e \in \mathbb{N}$, $f = \phi_e^{(n)}$ which is the same equality as that in (1), written with no reference to the "pairs" $(\vec{x}_n, y)$ that may or may not belong to the set of input-output pairs that a function such as f is.

Alternative names for "partial computable" and "computable" are *partial recursive* and *recursive* respectively.

We denote the set of all partial recursive functions by $\mathcal{P}$. Its subset of total recursive functions is denoted by $\mathcal{R}$. □

THEOREM B.1.2 (Smn or parametrisation theorem). *There is a* total *computable function* S_1^n[36] *of* $n+1$ *arguments such that the following holds for all* $e, z, \vec{x}_n$

$$\phi_e^{(n+1)}(z, \vec{x}_n) = \phi^{(n)}_{S_1^n(e,z)}(\vec{x}_n)$$

The proof of the above is rather easy, see for example [40], and the intuition behind it is that we can "demote" any (input) variable, z, into a "parameter" status, by replacing the instruction "**read** z" by the instruction "$z \leftarrow a$."

Thus the old program e can be trivially transformed into a program where the value for z is not read, but is assigned (say, the value is a).

The infinite sequence of such programs —for $a = 0, 1, 2, 3, \ldots$— with parameter z is constructed by the function S_1^n from e and from the parameter z. The essence of the word "constructed" is that the function S_1^n ends up being computable.

Trivially, for each value read into the $\vec{x}_n$ and each value a *read* into z in the case of program e —not read but *assigned* to z in the program $S_1^n(e, a)$— the two programs either both decline an answer (are in an "infinite loop") are both give the same numerical answer.

In what follows let us adopt Church's "λ-notation".

DEFINITION B.1.3 (Church's λ notation). We write

$$\lambda \vec{x}_n.\text{expression} \tag{1}$$

to indicate a function that for each (vector) value inputted in the $\vec{x}_n$ yields an output depicted by "expression".[37]

Thus, "λ" and the following "." act as "begin-end" brackets and declare the *ordered list* of input variables. The "expression" that follows the "." indicates the "rule" that produces the output (if any).

This notation allows one to take care to distinguish a *function* $f = \lambda \vec{x}_n.expression$ from *a function call* to f with specific, or unspecified, inputs like $f(\vec{w}_n)$.

[36] In a careful exposition of the proof, it turns out that S_1^n is primitive recursive (B.1.6).

[37] "expression" may or may not "depend" on $\vec{x}_n$: E.g., $\lambda x.x + 5$ versus $\lambda x.8$

Some of the literature uses

$$\vec{x}_n \mapsto \text{expression}$$

in place of (1) above. □

One of the niceties of lambda notation is, for example, to easily write down without explanatory words, a function with "don't care" arguments (some call them "dummy" arguments) as in $\lambda xyz.y$. This function, for all input values x, y, z returns the value y.

We now state

THEOREM B.1.4 (The Universal function theorem). *There is an e such that*

$$\phi_e^{(n+1)} = \lambda z \vec{y}_n . \phi_z^{(n)}(\vec{y}_n)$$

The proof of the above is also very easy and, for example, if one chooses the URM as model of computation it can be found in [40].

By the above theorem and Definition B.1.1 we have that $\lambda z \vec{y}_n . \phi_z^{(n)}(\vec{y}_n)$ is partial recursive.

There are a number of convenient basic results that we list without proof (cf. [40]). First a definition.

DEFINITION B.1.5 (Substitution and Primitive Recursion).

- Given $\lambda \vec{x} z \vec{w} . f(\vec{x}, z, \vec{w})$ and $\lambda \vec{u} . g(\vec{u})$. We say that

  $$\lambda \vec{x} \vec{u} \vec{w} . f(\vec{x}, g(\vec{u}), \vec{w})$$

  is obtained from the given functions by *substitution*.

- Special cases of "substitution" ([19])

  1. (Substitution of a constant)
     From $\lambda \vec{u}_n . g(u_1, u_2, \ldots, u_i, \ldots, u_j, \ldots, u_n)$ obtain

     $$\lambda \vec{u}_n^{\widehat{i}} . g(u_1, u_2, \ldots, k, \ldots, u_j, \ldots, u_n)$$

     where "$\vec{u}_n^{\widehat{i}}$" means that from the vector $\vec{u}_n$ the entry u_i is missing.
  2. (Swapping of variables)
     From $\lambda \vec{u}_n . g(u_1, u_2, \ldots, u_i, \ldots, u_j, \ldots, u_n)$ obtain

     $$\lambda \vec{u}_n . g(u_1, u_2, \ldots, u_j, \ldots, u_i, \ldots, u_n)$$

3. (Identification of variables)
   From $\lambda\vec{u}_n.g(u_1, u_2, \ldots, u_i, \ldots, u_j, \ldots, u_n)$ obtain

$$\lambda\vec{u}_n^{\widehat{j}}.g(u_1, u_2, \ldots, u_i, \ldots, u_i, \ldots, u_n)$$

4. (Expansion of the variables list)
   From $\lambda\vec{u}_n.g(u_1, u_2, \ldots, u_i, \ldots, u_j, \ldots, u_n)$ obtain

$$\lambda\vec{x}\vec{u}_n.g(u_1, u_2, \ldots, u_i, \ldots, u_j, \ldots, u_n)$$

   The first bullet above and the operations 1–4 collectively are known as the *Grzegorczyk Operations* or as *Grzegorczyk Substitution.*

- Next, given $\lambda\vec{y}.h(\vec{y})$ and $\lambda x\vec{y}v.g(x, \vec{y}, v)$ we say that $\lambda x\vec{y}.f(x, \vec{y})$ is defined by h and g via primitive recursion iff for all $x, \vec{y}$ the following two equalities hold:

$$f(0, \vec{y}) = h(\vec{y})$$
$$f(x + 1, \vec{y}) = g(x, \vec{y}, f(x, \vec{y}))$$

  We write $f = prim(h, g)$. □

It is easy —essentially by "programming" in the adopted model of computation— to show that $\mathcal{P}$ and $\mathcal{R}$ are, both, *closed* under Grzegorczyk substitution and primitive recursion, meaning that if the two functions that are given in the first bullet are in $\mathcal{P}$ then so is the resulting function under substitution and similarly for $f = prim(h, g)$:

Indeed, for the case of Grzegorczyk operations this is easy. For the first bullet (the general case), given the inputs, we compute the inner call first (to $g(\vec{u})$) and, once we get the result a, we then call $f(\vec{x}, a, \vec{w})$.

Of course $g(\vec{u})$ might be undefined, in which case so is the result of the substitution for the given inputs. Even if a *was* obtained by the first call, the second call —with input $\vec{x}, a, \vec{w}$— might fail.

The case for 1–4 above is trivial.

Clearly, if f and g are in $\mathcal{R}$, then they are total and thus so is the resulting $\lambda\vec{x}\vec{u}\vec{w}.f(\vec{x}, g(\vec{u}), \vec{w})$ since neither the first nor the second call can fail. A more trivial remark applies to cases 1–4.

As for primitive recursion, it is trivial to see that the pseudo code below computes $f(x, \vec{y})$ for any $x, \vec{y}$. If the model of computation is *adequate* (TM and URM are), then the pseudo code can be translated into, say, URM code and thus offer a definitive proof.

$$\begin{array}{ll} i & \leftarrow 0 \\ z & \leftarrow h(\vec{y}) \\ \textbf{loop} & x \textbf{ times} \\ z & \leftarrow g(i, \vec{y}, z) \\ i & \leftarrow i + 1 \\ \textbf{endloop} & \end{array}$$

Thus $\mathcal{P}$ is closed under primitive recursion, and so is $\mathcal{R}$ since an easy induction on x (with $\vec{y}$ a parameter) shows that $f = prim(h, g)$ is total if h and g are.

The operations in B.1.5 are ubiquitous in computability and using them one defines a set of *primitive recursive functions*, $\mathcal{PR}$, that turns out to be a proper subset (cf. [40]) of $\mathcal{R}$.

DEFINITION B.1.6 (Primitive recursive functions). $\mathcal{PR}$ is the set (computability theorists often say "class") of functions each of which is obtained by a *finite* application of "prim" and Grzegorczyk substitutions (B.1.5) repeatedly applied to earlier obtained functions staring from the *initial functions* $S = \lambda x.x + 1$, $\lambda x.k$ and $\lambda x.x$. □

Gödel used exclusively functions from $\mathcal{PR}$ in his Incompleteness theorems ([14]). These functions are originally due to Dedekind ([8, 9]).

It is convenient (but not theoretically necessary) to extend the notion of "computable" to relations on numbers from $\mathbb{N}$. This is done by coding the "outputs" —true, false— of relations by the numbers 0 and 1 respectively turning them into functions $\mathbb{N}^n \to \{0, 1\}$. Actually, we do it more subtly, not changing the ontology of relations: we use their *characteristic functions* as proxies.

DEFINITION B.1.7. Given a relation $R(\vec{x})$, its *characteristic function* c_R is given, for all $\vec{x}$, by

$$c_R(\vec{x}) \overset{Def}{=} \begin{cases} 0 & \text{if } R(\vec{x}) \text{ is true} \\ 1 & \text{if } R(\vec{x}) \text{ is false} \end{cases}$$

A relation $R(\vec{x})$ is *(primitive) recursive* iff its characteristic function is.

A relation $R(\vec{u})$ is *semi-recursive* (alternatively called *recursively enumerable* (r.e.) or Σ_1) iff it is equivalent to $(\exists y)Q(y,\vec{u})$ for some primitive recursive $Q(y,\vec{u})$.

The symbols $\mathcal{PR}_*, \mathcal{R}_*$ and $\mathcal{P}_*$ denote respectively the names of the sets of the primitive recursive, recursive and Σ_1 relations. □

Note. Convention has it to not use λ notation on relations unless we want to derive some special variant from one: Examples

- From $Q(x,y,z)$ to obtain $\lambda xz.Q(x,y,z)$ (y is a parameter).
- From $Q(x,y,z)$ to obtain $\lambda xyzwu.Q(x,y,z)$.
- From $Q(x,y,z)$ to obtain $\lambda xyz.Q(z,y,x)$.
- No λ indicates that $Q(x,y,z)$ means $\lambda xyz.Q(x,y,z)$.

THEOREM B.1.8 (Closure properties of $\mathcal{PR}_*, \mathcal{R}_*$ and $\mathcal{P}_*$).

1. *$\mathcal{PR}_*$ and $\mathcal{R}_*$ are closed under Boolean operations and bounded quantification, $(\forall x < z)Q(x,\vec{u})$ and $(\exists x < z)Q(x,\vec{u})$.*
2. *$\mathcal{P}_*$ is closed under the positive Boolean operations $\wedge$ and $\vee$ as well as under bounded quantification, $(\forall x < z)Q(x,\vec{u})$ and $(\exists x < z)Q(x,\vec{u})$ and unbounded existential quantification, $(\exists x)$ $Q(x,\vec{u})$.*

DEFINITION B.1.9 (Definition by cases). If f below —defined from the f_i and R_j, for all $\vec{x}$— is a function, then we say that the schema is a *definition by cases.*

$$f(\vec{x}) = \begin{cases} f_1(\vec{x}) & \text{if } R_1(\vec{x}) \\ f_2(\vec{x}) & \text{if } R_2(\vec{x}) \\ \vdots & \vdots \\ f_k(\vec{x}) & \text{if } R_k(\vec{x}) \\ f_{k+1}(\vec{x}) & \text{otherwise} \end{cases}$$

□

It is an easy fact in computability that if the R_j are recursive, then f is recursive or partial recursive according as the f_i are recursive or partial recursive.

Similarly if the R_j and the f_i are *primitive* recursive, then so is f.

These easy results all hinge on the preceding definitions and the fact that the function

$$sw = \lambda xyz.if\ x = 0\ then\ y\ else\ z$$

is primitive recursive.[38]

There is a plethora of primitive recursive functions in common use in logic and computability as indicated in the example below.

Example B.1.10. The following functions and relations are primitive recursive.

1. (Proper subtraction) $\lambda xy.x \dot{-} y$ where

$$x \dot{-} y = \begin{cases} 0 & \text{if } x < y \\ x - y & \text{otherwise} \end{cases}$$

2. If $\lambda x\vec{y}.f(x,\vec{y}) \in \mathcal{PR}$ then so are

$$\lambda z\vec{y}.\sum_{i<z} f(i,\vec{y})$$

and

$$\lambda z\vec{y}.\prod_{i<z} f(i,\vec{y})$$

By standard convention, the empty sum ($z = 0$) is defined to be 0 and the empty product is defined to be 1. Thus $\sum_{i<z}$ and $\prod_{i<z}$ are *primitive recursive operations*.

Note that a non-strict bound "$i \leq z$" is the same as "$i < z + 1$" thus $\sum_{i\leq z}$ and $\prod_{i\leq z}$ are also primitive recursive operations by Grzegorczyk substitution.

3. $\lambda xy.x^y$. This function is redefined at $(x,y) = (0,0)$ to return 1.

4. $\lambda xy.\lfloor x/y \rfloor$. This function is redefined at $y = 0$ to return $x + 1$ for all x.

[38] Write h for $\lambda yz.y$ and g for $\lambda xyzw.z$. These are in $\mathcal{PR}$ by Grzegorczyk operations starting from $\lambda x.x$. Then $sw(0,y,z) = h(y,z)$ and $sw(x+1,y,z) = g(x,y,z,sw(x,y,z))$.

5. $\lambda xy.rem(x, y)$. This function is redefined at $y = 0$ to return x for all x.
6. $x \mid y$. It says "x divides y (with 0 remainder)".
7. $Pr(x)$. It says "x is a prime".
8. $\lambda n.p_n$. The "n-th prime function".
9. $Seq(z)$. It says that "$z > 1$ and its prime number factorisation has *no missing primes* between 2 and the largest prime factor of z".
10. $\lambda x.lh(x)$. If $Seq(x)$ holds then this function returns the number of distinct primes n in the prime factorisation of $x = p_0^{a_0} p_1^{a_1} \cdots p_{n-1}^{a_{n-1}}$.
11. $\lambda nx.exp(n, x)$. This is the exponent of p_n in the prime factorisation of x.
12. $\lambda nx.(x)_n \stackrel{Def}{=} exp(n, x) \dot{-} 1$.

□

DEFINITION B.1.11 (Coding of sequences). A sequence of numbers,

$$a_0, a_1, \ldots, a_{n-1}$$

is coded by

$$\prod_{i<n} p_i^{a_i+1} \tag{1}$$

The code (1) is denoted by $\langle a_0, a_1, \ldots, a_{n-1} \rangle$.

Because of the "+1" in the exponents, we have that $Seq(\langle a_0, a_1, \ldots, a_{n-1} \rangle)$ is true. Thus $lh(\langle a_0, a_1, \ldots, a_{n-1} \rangle) = n$.

We "decode" by noting that $\Big(\langle a_0, a_1, \ldots, a_{n-1} \rangle \Big)_i = a_i$, for $i < n$. For $i \geq n$ we get the "don't care" $\Big(\langle a_0, a_1, \ldots, a_{n-1} \rangle \Big)_i = 0$, but lh will readily identify what we do *not* care about.

Concatenation: We note that if $z = \langle a_0, a_1, \ldots, a_{n-1} \rangle$ then

$$z \prod_{i<m} p_{i+n}^{b_i+1} \tag{2}$$

is the code of the concatenation $(*)$ of the two sequences

$$\langle a_0, a_1, \ldots, a_{n-1} \rangle * \langle b_0, b_1, \ldots, b_{m-1} \rangle$$

The expression in (2) more generally is a call to the primitive recursive function

$$\lambda zw.z \prod_{i<lh(w)} p_{i+lh(z)}^{b_i+1} \tag{3}$$

that we denote simply $\lambda zw.z * w$. □

Theoretical computation models like URMs and their computation can be arithmetized by the same technique invented by Gödel in the formal arithmetic and its proofs, by the tools in B.1.11.[39] Thus we have the following definition.

DEFINITION B.1.12 (The Kleene T-predicate). For each $n \geq 1$, the Kleene predicate[40] $T^{(n)}(z, \vec{x}_n, y)$ holds iff the n-input URM coded by z with input $\vec{x_n}$ has a terminating computation coded by y. □

One can prove (e.g., [40])

THEOREM B.1.13. *$T^{(n)}(z, \vec{x_n}, y)$ is primitive recursive.*

COROLLARY B.1.14 (The Kleene Normal Form theorem).

1. For any URM M of code z and n inputs, we have that M is defined on the input $\vec{x}_n$ iff $(\exists y)T^{(n)}(z, \vec{x}_n, y)$.
2. There is a primitive recursive function d such that for any $\lambda\vec{x}_n.f(\vec{x}_n) \in \mathcal{P}$ there is a number z and we have for all $\vec{x}_n$

$$f(\vec{x}_n) = [41]d((\mu y)T^{(n)}(z, \vec{x}_n, y))$$

The Kleene normal form theorem is essentially an elaboration of the statement in B.1.4, that is,

$$\phi_z^{(n)} = \lambda\vec{x}_n.d((\mu y)T^{(n)}(z, \vec{x}_n, y))$$

We usually write ϕ_z for $\phi_z^{(1)}$ and $T(z, x, y)$ for $T^{(1)}(z, x, y)$.

[39] Gödel actually carried out his arithmetisation with a coding devised via his "β function" (cf. [34,38])

[40] Strictly speaking, Kleene "relation". Computability practitioners often call their relations "predicates" (by abuse of logic terminology). This is one such instance.

[41] Equality is extended to hold when both sides are undefined.

B.2. The recursion theorem

THEOREM B.2.1 (Kleene's Recursion Theorem). *If $\lambda z \vec{x}_n . f(z, \vec{x}_n) \in \mathcal{P}$, then for some index e, we have $\phi_e(\vec{x}_n) \simeq f(e, \vec{x}_n)$ for all $\vec{x}_n$.*

PROOF: By assumption on f and Grzegorczyk operations,

$$\lambda z \vec{x_n} . f(S_1^n(z, z), \vec{x_n}) \in \mathcal{P}$$

Thus, for some $a \in \mathbb{N}$, we have $\phi_a^{(n+1)} = \lambda z \vec{x}_n . f(S_1^n(z, z), \vec{x}_n)$. Then by B.1.2,

$$\begin{aligned} f(S_1^n(a, a), \vec{x}_n) &= \phi_a^{(n+1)}(a, \vec{x}_n) \\ &= \phi_{S_1^n(a,a)}^{(n)}(\vec{x}_n) \end{aligned}$$

Take $e = S_1^n(a, a)$. □

B.3. Gödel's Diagonalisation lemma

The Gödel ([5, 14]) "*substitution function*" and diagonalisation or fixpoint lemma ("diagonal lemma" in [4]) is a precursor of Kleene's Smn function and recursion theorem respectively. The former are in the domain of (formal) Peano Arithmetic (PA) and were the tools through which Gödel formulated —in the language of PA— a sentence that says "I am not a (PA-)theorem". This sentence is the famous Gödel "*undecidable sentence*" in the sense that PA proves neither it nor its negation.

Just as the Smn function(s) and the recursion theorem of computability are products of arithmetisation, so are the substitution function and diagonalisation lemma of logic.

We are not going to conduct any arithmetisation in detail here (the interested reader can consult [21, 38]), but we will provide, we hope, enough insight as to how things work. Needless to say that the exact version of arithmatisation is not important, thus we will pretend that prime-power arithmetisation is what took place.

A version of this kind of arithmetisation —but as applied to Turing Machines— can be found in Davis ([7]).

In brief, let us fix the order of elements in the PA alphabet,

$$x, p, \#, =, (,), \bot, \neg, \wedge, \forall; 0, S, +, \times, < \tag{1}$$

where the purpose of all symbols is evident, with the possible exception of "$x, p, \#$".

These are used to generate *object* variables,

$$(x\#), (x\#\#), (x\#\#\#), \ldots$$

which we will metatheoretically write down as

$$v_i \text{ for the actual } (x\overbrace{\#\#\ldots\#}^{i+1})$$

thus v_0 stands for $(x\#)$.

One more relaxation of notation —leading to more convenience— will have the bold face $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{v}, \mathbf{w}$, with or without primes or subscripts, stand for any v_i.

Similarly, p and $\#$ generate all Boolean variables,

$$(p\#), (p\#\#), (p\#\#\#), \ldots$$

which we will metatheoretically write down as

$$p_i \text{ for the actual } (p\overbrace{\#\#\ldots\#}^{i+1})$$

thus p_0 stands for $(p\#)$.

One more relaxation of notation —leading to more convenience— will have the bold face $\mathbf{p}, \mathbf{q}, \mathbf{r}$, with or without primes or subscripts, stand for any p_i.

Now assign the first 15 odd numbers, in order, to the symbols in (1) in the order given:

x	p	$\#$	$=$	$($	$)$	$\bot$	$\neg$	$\wedge$	$\forall$	0	S	$+$	$\times$	$<$
1	3	5	7	9	11	13	15	17	19	21	23	25	27	29

Notation. **Informally**, let us indicate that a PA-alphabet *symbol* s has numerical code a by writing

$$gn\{s\} = a \tag{2}$$

Next, if $A = a_1 a_2 \ldots a_n$ is a string over the alphabet (1) we code it as

$$gn\{A\} \stackrel{Def}{=} \langle gn\{a_1\}, gn\{a_2\}, \ldots, gn\{a_n\}\rangle \tag{3}$$

Finally, if $\mathcal{A} = A_1, A_2, \ldots, A_n$ is a sequence of strings over the alphabet (1) —for example a proof— then we code it as

$$gn\{\mathcal{A}\} \stackrel{Def}{=} \langle gn\{A_1\}, gn\{A_2\}, \ldots, gn\{A_n\}\rangle \tag{4}$$

Remark B.3.1. It is important to note that

- The code of an alphabet symbol is odd.
- The code of a string over the alphabet is even and every prime factor has an even exponent (odd plus 1).
- The code of a sequence of strings over the alphabet is even and every prime factor has an odd exponent (even plus 1).

This is a good start towards encoding PA, but is all happening in the *metatheory* of PA as the codes are natural numbers. One last step in order to formalise (2)–(4) is to

1. use numerals, the PA counterparts for the informal numbers. The numeral for 0, denoted $\widetilde{0}$ is the 0 from (1), while the numeral for $n+1$ —denoted $\widetilde{n+1}$— is
$$\overbrace{SS\ldots S}^{n+1\ copies}0$$
2. To use the fact ([21, 34, 38]) that we can define every $\mathcal{PR}$ function *within* PA, that is formally.

Showing how this is done is extremely tedious (see the accounts in loc. cit.) and beyond the scope of this thesis. Suffice it to say that informal definitions of primitive recursive functions can be mirrored inside PA —that is formally— and thus we we obtain the coding (2)–(4) inside PA.

The traditionally used symbol for formal Gödel numbers (codes) is $\ulcorner \ldots \urcorner$.

Thus in formal notation, and with the understanding that the coding has been done within PA, (2)–(4) translate as

$$\ulcorner s \urcorner = \widetilde{a} \tag{2'}$$

$$\ulcorner A \urcorner \stackrel{Def}{=} \langle \ulcorner a_1 \urcorner, \ulcorner a_2 \urcorner, \ldots, \ulcorner a_n \urcorner \rangle \tag{3'}$$

and

$$\ulcorner \mathcal{A} \urcorner \stackrel{Def}{=} \langle \ulcorner A_1 \urcorner, \ulcorner A_2 \urcorner, \ldots, \ulcorner A_n \urcorner \rangle \tag{4}$$

We have in general

$$n = gn\,\{\mathcal{E}\} \text{ iff } \vdash_{PA} \widetilde{n} = \ulcorner \mathcal{E} \urcorner$$

Another way to say the above is

$$\vdash_{PA} \ulcorner \mathcal{E} \urcorner = \widetilde{gn\,\{\mathcal{E}\}}$$

where the expression $\mathcal{E}$ might be an alphabet symbol, a string, or a finite sequence of strings.

THEOREM B.3.2 (Gödel's substitution theorem). *There is a formal (defined within PA) primitive recursive function "sub" of two arguments, such that if A is a formula where the variable v_0 occurs free —A may have other free variables as well— and $\vdash_{PA} x = \ulcorner A \urcorner$, then $sub(x, y)$ provably (in PA) returns $\ulcorner A[v_0 := y] \urcorner$, where y is some numeral $\widetilde{n}$.*

LEMMA B.3.3 (Fixed Point, or Fixpoint, or Diagonalisation Lemma). *For any formula A of Peano Arithmetic that has $v_0, \mathbf{x}_1 \ldots, \mathbf{x}_n$ as its free variables, there is a formula $B(\mathbf{x}_1 \ldots, \mathbf{x}_n)$ such that $\vdash_{PA}$ $B(\mathbf{x}_1 \ldots, \mathbf{x}_n) \leftrightarrow A(\ulcorner B(\mathbf{x}_1 \ldots, \mathbf{x}_n) \urcorner, \mathbf{x}_1 \ldots, \mathbf{x}_n)$.*

PROOF: Throughout the proof *sub* is the formal substitution function from B.3.2.

For some $d \in \mathbb{N}$, PA proves $\widetilde{d} = \ulcorner A(sub(v_0, v_0), \mathbf{x}_1 \ldots, \mathbf{x}_n) \urcorner$. Now, $\ulcorner A(sub(\widetilde{d}, \widetilde{d}), \mathbf{x}_1 \ldots, \mathbf{x}_n) \urcorner$ is provably (in PA) equal to $sub(\widetilde{d}, \widetilde{d})$, by B.3.2. Thus, if we set $\widetilde{e} = sub(\widetilde{d}, \widetilde{d})$ we obtain

$$\widetilde{e} = \ulcorner A(\widetilde{e}, \mathbf{x}_1 \ldots, \mathbf{x}_n) \urcorner [42] \qquad (i)$$

From

$$A(\widetilde{e}, \mathbf{x}_1 \ldots, \mathbf{x}_n) \leftrightarrow A(\widetilde{e}, \mathbf{x}_1 \ldots, \mathbf{x}_n)$$

we obtain —using (i) only on the right hand side of $\leftrightarrow$,

$$A(\widetilde{e}, \mathbf{x}_1 \ldots, \mathbf{x}_n) \leftrightarrow A(\ulcorner A(\widetilde{e}, \mathbf{x}_1 \ldots, \mathbf{x}_n) \urcorner, \mathbf{x}_1 \ldots, \mathbf{x}_n)$$

$A(\widetilde{e}, \mathbf{x}_1 \ldots, \mathbf{x}_n)$ is the "B" we want. □

[42] Up to this point the proof is almost identical to that of B.2.1.

Appendix C

Solovay's Theorem

C.1. Introduction

Solovay's theorem [37] states the arithmetical completeness of GL, hence establishing that this propositional modal logic models (meaning here "imitates") *provability in PA*.

C.2. The Logic GL

GL is the propositional fragment of ML^3, that is, its language omits all first-order symbols, and the GL axioms are those of ML^3 minus the first-order axioms.

C.2.1. Axioms

DEFINITION C.2.1 (Basic Axiom Schemata of GL).

GL1 All tautologies

GL2 $\Box(A \to B) \to \Box A \to \Box B$

GL3 $\Box(\Box A \to A) \to \Box A$

GL4 $\Box A \to \Box\Box A$

We will denote by Λ —with apologies to the reader for the possible but unlikely confusion with the "Λ" of ML^3— the set of *all instances* of the schemata GL1 – GL4.

The set of *axioms* is $\Lambda \cup \Box\Lambda$.

As we commented in the case of ML^3, the schema GL4 is derivable from the other axioms, but in the interest of making necessitation a weak *admissible* (rather than a strong *primary*) rule the inclusion of GL4 allows us to avoid taking $\Lambda \cup \Box\Lambda \cup \Box\Box\Lambda$ as the set of axioms. □

C.2.2. Rules of Inference

DEFINITION C.2.2. GL has only *one* primary rule, modus ponens (MP), and we note that a proof like that for 4.3.6 shows that *weak necessitation* is an admissible rule.) □

C.3. Kripke Semantics

The Kripke frames and structures used here have essentially the same definitions as 5.0.1 and 5.0.2, minus first-order clauses.

As such we omit the definitions since they adapt to GL. Everything we did under the above heading in the main body of this thesis is trivially adapted from the ML^3 case to the GL case by simply omitting first-order cases in our definitions and proofs.

C.4. Semantic Completeness

The proof of completeness of GL with respect to *finite* reverse well-founded Kripke models follows closely from the completeness of its subtheory K4 obtained by removing Łöb's axiom (GL3 above). This is entirely analogous to the approach we took in passing from M^3 models to *finite* reverse well-founded ones for ML^3, thus as the propositional proofs are, essentially, identical to the first-order ones but with the first-order concepts removed, we will not repeat such propositional variants here. We will state then, without proof,

THEOREM C.4.1. *GL is complete for finite, irreflexive and transitive Kirpke models.*

Remark C.4.2. How exactly one goes about constructing a *finite reverse well-founded* model for GL (or for that matter for ML^3) admits quite a bit of variation in the literature and it so happens that Solovay's proof of his arithmetical completeness theorem is indifferent to both the version of the construction and to the ontology of the worlds that such a construction entails.

C.5. Arithmetical Completeness

THEOREM C.5.1. *GL is arithmetically complete in PA in the sense that, for any A over the language of GL, if all arithmetical realisations A^* of A are provable in PA, then A is provable in GL.*

The above has a very short proof that is entirely analogous to that for ML^3. Cf. the two lines of conclusion on p.49.

The burden of proof is on

- Solovay's Lemmata (cf. 7.2.3 and also below C.5.2) —which, luckily, as we noted already (cf. p.43, under the construction of $\widehat{R}^0$) are readily applicable *as is* without having to refer to their complex proofs,

  and

- the proof of "the main lemma" —which we omit as it is quite similar and quite easier than the case for ML^3 (cf. 7.2.10) as the absence of first-order structure helps.

Notwithstanding the assertion in the first bullet above, in the interest of a panoramic view of what is at play in this thesis we give reasonably complete proofs of the Solovay Lemmata here.

Let then $\langle W, R\rangle$ be a finite frame (cf. 5.0.1), where

$$W = \{0, 1, 2, 3, ..., n\}$$

and R be irreflexive and transitive, thus automatically reverse well-founded —else we would have cycles (by finiteness)

$$\underline{i_1} R i_2 R i_1 R i_3 \ldots i_m R \underline{i_1}$$

in the graph structure, and hence we would have $i_1 R i_1$ by transitivity thus contradicting irreflexivity.

Solovay embeds the graph (W, R) in PA (or any consistent recursive extension $\mathcal{T}$[43]) and proves

LEMMA C.5.2 (Solovay's Lemmata). *$\mathcal{T}$ is some consistent recursive extension of PA over a finite extension of the PA language. There are sentences S_i, for $0 \leq i \leq n$, of the language, such that*

[43] That is, it extends PA and its axiom set is recursive.

(1) *For all* $i \neq j$, $\vdash_{\mathcal{T}} \neg S_i \vee \neg S_j$.

(2) *For* $0 \leq i \leq n$, $\mathcal{T} + S_i$ *is consistent.*

(3) *If* iRj, *then* $\vdash_{\mathcal{T}} S_i \rightarrow \neg\Theta(\ulcorner \neg S_j \urcorner)$, *where* Θ *is the provability predicate for* $\mathcal{T}$.

(4) *If* $i > 0$, *then* $\vdash_{\mathcal{T}} S_i \rightarrow \Theta(\ulcorner \bigvee_{iRj} S_j \urcorner)$.

([35, 37]) The proof starts by defining a function $h : \mathbb{N} \rightarrow W$ and proving *within* $\mathcal{T}$ the properties of its "limit l", which are stated as properties (1)–(4) in C.5.2.

We say that a function with inputs from $\mathbb{N}$ and outputs also in $\mathbb{N}$ has a limit l iff, for some m, $x \geq m$ implies $h(x) = l$. The limit manifests its presence in the lemmata above via the sentences S_i where "S_i" is an abbreviation of $\mathbf{l} = \widetilde{i}$, that is, a statement which when stated informally says that the limit l is equal to $i \in W$. In "$\mathbf{l} = \widetilde{i}$" "$\mathbf{l}$" is a closed term[44]and "$\widetilde{i}$" is introduced on p.68.

Solovay defined his "h" by a recursion where h "walks" on the nodes of the graph[45] by following the edges. Thus for any $i \in W$, if $h(x) = i$ for some x, then $h(i+1) = i$ or $h(i+1) = j$, where iRj holds. The finiteness of the graph and the irreflexivity guarantee intuitively that a limit exists, since one must have eventually (for a "large" x and then onwards) all output values form a constant sequence, since the alternative is to have a cycle (a node is visited twice, which negates irreflexivity).

Letting on a bit more detail before we get technical, the exact definition of h makes it hesitant to progress (its outputs) along the graph as it will set $h(i + 1) = j$, where iRj (of course $i \neq j$ by irreflexivity) holds only if it (the definition) "knows" that PA can *prove* —indeed, with a proof of Gödel number $x + 1$— that j is *not* the limit.

Our exposition is influenced by [4, 10, 35, 37]. The last source in particular has an uncomplicated way to define h primitive recursively that makes the exposition (and the technical details) more passable.

[44]We will never have to show this term explicitly as it *always* enters in our discussions implicitly in sentences like $\mathbf{l} = \widetilde{i}$ that we call simply S_i.

[45]The graph has as nodes the members of W. An edge exists from i to j iff iRj.

The first approximation to the *informal* (outside PA, that is) definition of h is ([35,37]) in terms of the limit $l = \lim_{x\to\infty} h(x)$ of h, that is, $(\exists y)(\forall x)(x \geq y \to h(x) = l)$. Of course, we do not have l in closed form (nor do we know that it exists *at the time we are attempting the definition*).

$$\begin{array}{l} h(0) = 0 \\ h(x+1) = j \text{ if } h(x)Rj \wedge Proof(x+1, gn\,\{l \neq j\}) \\ h(x+1) = h(x) \text{ otherwise} \end{array}$$

Above we wrote "$gn\,\{l \neq j\}$" rather than "$\ulcorner l \neq j \urcorner$" since the above approximation to the definition is informal (outside PA). $Proof(u, v)$, as before (cf. p.13), is the *primitive recursive* relation "u codes a PA (or a $\mathcal{T}$) proof and v occurs as (the Gödel number of) a formula in said proof".

Remark C.5.3. In general, we can only assert that $Proof(u, v)$ is *recursive* (but not *primitive* recursive) as it inherits this property from the set of axioms. However, in this thesis we deal with PA, which has a *primitive recursive* set of axioms ([21,38]) and so does our "$\mathcal{T}$" that extends PA by adding *finitely many* axioms (cf. p.45). □

This definition of h can be shown —using the recursion theorem in the metatheory— to be recursive, once we express "$gn\,\{l \neq j\}$" via h, but we will do better than this and will present a closely related definition of h that is obviously primitive recursive. *We also need to say how we pick j.*

The authors of [10] add an argument to h and also, as in [4], they replace the condition "$Proof(x+1, gn\,\{l \neq j\})$" above with a more manageable one:

Definition C.5.4. We define informally a *primitive recursive* function $\lambda uv.notlim(u, v)$ so that *if we input into u the Gödel number of a formula $F(x, y)$* that represents the graph $f(x) = y$ of a function, and *if we also input the Gödel number of a numeral $\widetilde{k}$* into v it returns the Gödel number of the expression

$$\neg(\exists w)(\forall x)\Big(x \geq w \to F(x, v)\Big) \tag{1}$$

which says that v is *not* the limit of the function f whose graph is $F(x, y)$. □

Using the "Tarski" (cf. [38]) trick that avoids using "sub" in expressing substitutions —and this in turn rests on the so called "one-point-rule"[46] from York University's MATH 1090 (cf. [39])— we can rewrite (1) above as

$$\neg(\exists w)(\forall x)\Big(x \geq w \to (\exists y)\big(y = v \wedge F(x,y)\big)\Big) \tag{2}$$

Definition C.5.4 requires the output of *notlim* to be the Gödel number of the entire formula in (2). To prepare for this we make two observations,

1. Since the construction of the Gödel code of a string or a finite sequence of strings proceeds from left to right we have (informally) $gn\{AB\} = gn\{A\} * gn\{B\}$ and (formally in PA, or a *recursive extension*[47] $\mathcal{T}$ thereof) $\ulcorner AB \urcorner = \ulcorner A \urcorner * \ulcorner B \urcorner$. This $*$ defined primitive recursively in Definition B.1.11 —not to be confused with the *raised* * of arithmetical realisations— is associative, since concatenation is, and thus in a chain like $\ulcorner A \urcorner * \ulcorner B \urcorner * \ulcorner C \urcorner$ we use brackets only if we absolutely must for clarity.

2. The requirement in the definition, "input the Gödel number of a numeral $\widetilde{k}$ into v", is easily accommodated by using the (formal, in PA or $\mathcal{T}$) primitive recursive function "Num" (e.g., [38]) defined as

$$\begin{aligned} Num(0) &= \ulcorner 0 \urcorner \\ Num(Sx) &= \ulcorner S \urcorner * Num(x) \end{aligned}$$

   An easy informal induction on n (outside PA) shows that for all n $Num(\widetilde{n}) = \ulcorner \widetilde{n} \urcorner$, thus, formally, for every numeral $\widetilde{n}$, $\vdash_{\mathcal{T}} Num(\widetilde{n}) = \ulcorner \widetilde{n} \urcorner$.

   The induction: For $n = 0$ this is so by the *Basis* of the definition of Num and the fact that $\widetilde{0}$ is defined to be 0.

   Now using "=" conjunctionally, with reasons provided on each

[46] The two equivalences below are theorems of pure logic, and thus of any first-order use of logic in mathematics and elsewhere. $(\exists x)(x = t \wedge A[x]) \leftrightarrow A[t]$ and $(\forall x)(x = t \to A[x]) \leftrightarrow A[t]$, where the term t contains no (free) x.

[47] An extension that has a recursive set of axioms. Of course, our $\mathcal{T}$ does better, with a primitive recursive set of axioms.

line, each occurrence of = provably equates its two sides.

$$\begin{array}{lll} Num(\widetilde{n+1}) = Num(S\widetilde{n}) & & \text{definition of } \widetilde{\ } \\ = \ulcorner S \urcorner * Num(\widetilde{n}) & & \text{by the 2nd recurrence for } Num \\ = \ulcorner S \urcorner * \ulcorner \widetilde{n} \urcorner & & \text{by the obvious I.H.} \\ = \ulcorner S\widetilde{n} \urcorner & & \text{by observation 1. above} \\ = \ulcorner \widetilde{n+1} \urcorner & & \text{definition of } \widetilde{\ } \end{array}$$

We can now compute the (formal) Gödel number of (2) with parameters u and v as given in C.5.4. The formal (defined by the explicit term-definition below) *notlim* in PA or $\mathcal{T}$ is given below and as such it is clearly a (formal) primitive recursive function of u and v

$$\begin{array}{l} notlim(u,v) \overset{Def}{=} \\ \ulcorner \neg(\exists w)(\forall x)\Big(x \geq w \rightarrow (\exists y)\big(y = \urcorner * Num(v) * \ulcorner \wedge \urcorner * u * \ulcorner\big)\Big) \urcorner \end{array} \tag{3}$$

Thus, the argument u expects to receive an input such as $\ulcorner F(x,y) \urcorner$, for some formula $F(x,y)$ of PA (or $\mathcal{T}$), while v expects to receive a numeral $\widetilde{m}$, which Num converts into $\ulcorner \widetilde{m} \urcorner$, in which case (3) —for said inputs— trivially becomes (cf. observation 1. above):

$$\ulcorner \neg(\exists w)(\forall x)\Big(x \geq w \rightarrow (\exists y)\big(y = \widetilde{m} \wedge F(x,y)\big)\Big) \urcorner \tag{3'}$$

Remark C.5.5. Thus *notlim* says that $\widetilde{m}$ is *not* the limit of $f(x)$ whose graph is represented by the formula of PA ($\mathcal{T}$) $F(x,y)$. □

Next, we present the modified "h" from [10]. Let us call it $\widehat{h}$ to distinguish it from the "h" as introduced in [4, 35, 37]. $\widehat{h}$ has a parameter u and the recursion is through obvious calls to $\widehat{h}$ rather than via the obscure formula "$Proof(x+1, gn\,\{l = j\})$" that refers to the term l, the limit of h. The primitive recursion below is informal, and suppresses logical notation in favour of English text (e.g., the presence of j). We will formalise shortly.

$$\begin{array}{l} \widehat{h}(u,0) = 0 \\ \widehat{h}(u,x+1) = \text{ least } j \leq \widetilde{n} : \quad h(u,x)Rj \wedge Proof(x+1, notlim(u,j)) \\ \widehat{h}(u,x+1) = h(u,x) \text{ otherwise} \end{array} \tag{\dagger}$$

Remark C.5.6.

1. $\mathcal{PR}$ is closed under *bounded search* both metatheoretically and formally. For the formal case —which is based on the least

principle that is equivalent to induction— see [21, 34, 38].

For the informal case, one way to show closure under bounded search is as follows:

Given $\lambda y \vec{x}.f(y, \vec{x}) \in \mathcal{PR}$. Define

$$(\mu y < z) f(y, \vec{x}) \stackrel{Def}{=} \begin{cases} \min\{y : f(y, \vec{x}) = 0\} & \text{if } (\exists y < z) f(y, \vec{x}) = 0 \\ z & \text{otherwise} \end{cases}$$

We have that

$$(\mu y < z) f(y, \vec{x}) = \sum_{i<z} \left(1 \dot{-} \left(1 \dot{-} \prod_{j \leq i} f(j, \vec{x})\right)\right) \tag{4}$$

$\lambda z \vec{x}.(\mu y < z) f(y, \vec{x}) \in \mathcal{PR}$ by (4) and B.1.10, case 2.

As usual, "$(\mu y \leq z)$" means "$(\mu y < z + 1)$".

Using the characteristic function of a relation in $\mathcal{PR}_*$ as a proxy we can derive primitive recursive functions by doing bounded search on such relations:

$$(\mu y < z) R(y, \vec{x}) \stackrel{Def}{=} (\mu y < z) c_R(y, \vec{x})$$

2. We note that the *finite* xRy has a formal counterpart $\langle x, y \rangle \in \mathbf{R}$, where $\mathbf{R}$ is a formal representation of R in this sense:

   We view R as a finite sequence of coded pairs $\langle x, y \rangle$ such that xRy arranged in, say, ascending order. Thus, if the so ordered $\langle x, y \rangle$ result in the sequence

   $$r_0, r_1, r_2, \ldots, r_{k-1}$$

   then we define within PA ($\mathcal{T}$)

   $$\mathbf{R} \stackrel{Def}{=} \langle r_0, r_1, r_2, \ldots, r_{k-1} \rangle$$

   Thus xRy is true iff PA ($\mathcal{T}$) proves

   $$\langle x, y \rangle \in \mathbf{R}$$

where the primitive recursive relation $z \in w$ is defined in PA primitive recursively as

$$Seq(w) \wedge (\exists k < lh(w))(w)_k = z$$

In all formal contexts that follow we will write "$x\mathbf{R}y$" for "$\langle x, y\rangle \in \mathbf{R}$". □

As in 1.7.0.2 in [40], we will rewrite the informal definition of $\widehat{h}$, this time *formally*, by defining its *graph* —$\widehat{h}(u, x) = y$— by the *formula* $\widehat{H}(u, x, y)$ of $\mathcal{T}$ (PA).

DEFINITION C.5.7 (The $\widehat{H}(u, x, y)$). The relation $\widehat{H}(u, x, y)$ is introduced by the formal (primitive) recursive definition below.

$$\begin{aligned}
&\widehat{H}(u,x,y) \overset{Def}{\longleftrightarrow} \\
&\quad x = 0 \wedge y = 0 \\
&\vee \\
&\quad \Big(x > 0 \wedge y = (\mu z \leq x)(\exists w \leq \widetilde{n})\Big(Proof(x, notlim(u, z)) \wedge \\
&\qquad\qquad \big(\widehat{H}(u, x \dot{-} \widetilde{1}, w) \wedge w\mathbf{R}z\big)\Big)\Big) \\
&\vee \\
&\quad \Big(x > 0 \wedge \neg(\exists v \leq x)(\exists w \leq \widetilde{n})\Big(Proof(x, notlim(u, v)) \wedge \\
&\qquad\qquad w\mathbf{R}v \wedge \widehat{H}(u, x \dot{-} \widetilde{1}, w)\Big) \wedge \widehat{H}(u, x \dot{-} \widetilde{1}, y)\Big)
\end{aligned}$$

□

Remark C.5.8.

1. As $v \leq x$ and $z \leq x$ since x is a Gödel number of a proof where z or v occur, the bounded search "$(\mu z \leq x)$" and the bounded quantifier "$(\exists v \leq x)$" miss no z that we want to compare with w in $w\mathbf{R}z$ in the disjunct #2 above, or v that we want to compare with w in $w\mathbf{R}v$ in the disjunct #3 above.

2. By the informal 1.7.0.2 in [40] —formalised accounts of primitive recursive definitions occur in [21, 34, 38]— a unique *primitive recursive* formula ("unique" within provable equivalence) $\widehat{H}(u, x, y)$ satisfying the formal recurrence above exists. It is moreover easy to see, at least informally, that it is the representing formula for $\widehat{h}$ which is given by the same recurrence (†)

although the latter is formulated "functionally" and informally on p.76. As we will exclusively use the fixpoint $H(x,y)$ below towards proving C.5.2 we will not pursue this verification nor will we promote $\widehat{h}$ from its motivational role to a formal role.

Now, a formal induction on x within PA (or $\mathcal{T}$) quickly shows that $\widehat{H}(u,x,y)$ is *total* and *single-valued* in y (C.5.9 below).

$$\vdash_{\mathcal{T}} (\forall x)(\exists! y)\widehat{H}(u,x,y)$$

where as is standardly the case "$\exists!$" states *unique existence*. See Lemma C.5.9 below. □

By the Diagonalisation Lemma (B.3.3) there is a formula $H(x,y)$ such that

$$\vdash_{\mathcal{T}} H(x,y) \leftrightarrow \widehat{H}(\ulcorner H(x,y) \urcorner, x, y)$$

thus, from the recurrence $\widehat{H}$ above, and using the substitution $u := \ulcorner H(x,y) \urcorner$, the $\widehat{H}$ recurrence produces the theorem (about the fixpoint H) below:

$$\begin{aligned}
&\vdash_{\mathcal{T}} H(x,y) \longleftrightarrow \\
&x = 0 \wedge y = 0 \\
\vee\ &\Big(x > 0 \wedge y = (\mu z \leq x)(\exists w \leq \widetilde{n})\Big(Proof(x, notlim(\ulcorner H(x,y) \urcorner, z)) \wedge \\
&\qquad \big(H(x \dot{-} \widetilde{1}, w) \wedge w \mathbf{R} z \big) \Big) \Big) \qquad (1) \\
\vee\ &\Big(x > 0 \wedge \neg(\exists v \leq x)(\exists w \leq \widetilde{n})\Big(Proof(x, notlim(\ulcorner H(x,y) \urcorner, v)) \wedge \\
&\qquad w \mathbf{R} v \wedge x \dot{-} \widetilde{1}, w) \Big) \wedge H(x \dot{-} \widetilde{1}, y) \Big)
\end{aligned}$$

LEMMA C.5.9. $\vdash_{\mathcal{T}} (\forall x)(\exists! y)\widehat{H}(u,x,y)$.

PROOF: Indeed

- (*Existence*) We examine each of the three definition clauses of $\widehat{H}$. The basis is the first clause and indeed a y exists ($= 0$) such that $\vdash_{\mathcal{T}} \widehat{H}(u,0,y)$.

  The two clauses for $x > 0$ being complementary, assuming one or the other leads to a y that works: the first on the I.H. that

a w exists such that $\vdash_{\mathcal{T}} \widehat{H}(u, x \dot{-} \widetilde{1}, w)$, the second on the I.H. that a y exists such that $\vdash_{\mathcal{T}} \widehat{H}(u, x \dot{-} \widetilde{1}, y)$.

- (*Uniqueness*) The Basis for $x = 0$ is trivial: $y = 0$.

  In clause #2, the y chosen using μ is unique since the w in $\widehat{H}(u, x \dot{-} \widetilde{1}, w)$ is by I.H.

  Lastly, in clause #3, the y chosen is unique by the I.H.

□

COROLLARY C.5.10. $\vdash_{\mathcal{T}} (\forall x)(\exists! y) H(x, y)$.

PROOF: By C.5.9, substituting the constant term $\ulcorner H(x,y) \urcorner$ into u. □

Thus $H(x, y)$ represents the graph of a total function.

Remark C.5.11. By (3′) preceding C.5.5, we have

$$\begin{aligned} &\vdash_{\mathcal{T}} notlim(\ulcorner H(x,y) \urcorner, \widetilde{m}) \leftrightarrow \\ &\quad \ulcorner \neg(\exists w)(\forall x)\Big(x \geq w \rightarrow (\exists y)\big(y = \widetilde{m} \wedge H(x,y)\big)\Big) \urcorner \end{aligned} \tag{2}$$

The right hand side of (2) is the Gödel number of the statement "$\neg \mathbf{S_m}$", that says

"The limit of $H(x, y)$ as x goes to infinity is *not* equal to $\widetilde{m}$". (3)

□

We need a few more lemmata before we prove C.5.2.

LEMMA C.5.12. $\vdash_{\mathcal{T}} H(x, y) \rightarrow y \leq \widetilde{n}$ *(recall: $W = \{0, 1, 2, \ldots, n\}$).*

PROOF: Formal induction on x:

- *Basis.* For $x = 0$ we have by equation (1) $\vdash H(0, 0)$. But PA proves $\vdash 0 \leq \widetilde{n}$ (7.1.1, axioms P7 and P9).

- Fix $x > 0$ and assume (I.H.) the statement for said x.

  Now, in clause #2 of (1) $\vdash y = z \wedge w\mathbf{R}z$ for some w. But R is on W (informally) thus (formally) $\vdash z = 0 \vee z = \widetilde{1} \vee \ldots \vee z = \widetilde{n}$. This says $\vdash z \leq \widetilde{n}$.

  In clause #3 of (1) we are done by the I.H. since we have $\vdash H(x \dot{-} \widetilde{1}, y)$.

□

LEMMA C.5.13. *We can prove in PA (or $\mathcal{T}$) that the y-limit of $H(x,y)$ as $x \to \infty$ is unique (if it exists).*

PROOF: Suppose we have

$$(\forall x)(x > u \to H(x,z)) \qquad (i)$$

and

$$(\forall x)(x > v \to H(x,w)) \qquad (ii)$$

We note without proof that $\vdash_{PA} u+v+\widetilde{1} > u$ and $\vdash_{PA} u+v+\widetilde{1} > v$ (see [38] how such simple facts can be proved from 7.1.1). Applying the specialisation rule of pure logic ("$(\forall z)A[z] \vdash A[t]$, for any term t") we get from (i) and (ii)

$$\vdash u+v+\widetilde{1} > u \to H(u+v+\widetilde{1},z)$$

and

$$\vdash u+v+\widetilde{1} > v \to H(u+v+\widetilde{1},w)$$

Thus

$$\vdash H(u+v+\widetilde{1},z) \qquad (i')$$

and

$$\vdash H(u+v+\widetilde{1},w) \qquad (ii')$$

by the inequalities we established above, and tautological implication. Now C.5.10 plus (i') and (ii') yield $z = w$. □

COROLLARY C.5.14 (Lemma C.5.2, (1)). For all i,j in W such that $i \neq j$ we have $\vdash_{\mathcal{T}} \neg(S_i \wedge S_j)$

PROOF: This is a rephrasing of C.5.13. □

LEMMA C.5.15.

$$\vdash_{\mathcal{T}} H(x,\widetilde{m}) \to (\forall z > x)\Big(H(z,\widetilde{m}) \vee \bigvee_{mRj} H(z,\widetilde{j})\Big) \qquad (\P)$$

Note. The notation "mRj", if we think of m as fixed, means that j is a "running index" such that mRj holds.

PROOF: We fix x and $\widetilde{m}$ and assume $H(x,\widetilde{m})$. That is, invoking the deduction theorem, we will work in $\mathcal{T}+H(x,\widetilde{m})$.

We prove by (formal) induction on the variable u

$$\vdash_{\mathcal{T}+H(x,\widetilde{m})} H(x+u+\widetilde{1},\widetilde{m}) \vee \bigvee_{mRj} H(x+u+\widetilde{1},\widetilde{j}) \qquad (\dagger)$$

Of course $\vdash_{\mathcal{T}} x+u+\widetilde{1} > x$ (cf. with the proof above).

- *Basis.* $u=0$. We want (omitting the subscript of $\vdash$)

  $$\vdash H(x+\widetilde{1},\widetilde{m}) \vee \bigvee_{mRj} H(x+\widetilde{1},\widetilde{j})$$

  This follows from (1) (p.79) since $\vdash H(x,\widetilde{m})$ we get either $\vdash H(x+\widetilde{1},\widetilde{m})$ or $\vdash H(x+\widetilde{1},\widetilde{k})$, for some k satisfying mRk.

- Assume now (I.H.) that we have (†) for a fixed u. We will establish then

  $$\vdash_{\mathcal{T}+H(x,\widetilde{m})} H(x+u+\widetilde{2},\widetilde{m}) \vee \bigvee_{mRj} H(x+u+\widetilde{2},\widetilde{j}) \qquad (\ddagger)$$

  to conclude the induction. We split the I.H. (†) into two cases (proof by cases).

Case (A) We add the hypothesis $H(x+u+\widetilde{1},\widetilde{m})$. By (1) on p.79 we have, just as in the *Basis* case, $\vdash H(x+u+\widetilde{2},\widetilde{m})$ or $\vdash H(x+u+\widetilde{2},\widetilde{k})$ for some k satisfying mRk.

Case (B) We add the hypotheses $H(x+u+\widetilde{1},\widetilde{j})$, for all j such that mRj. But then, as above we have, for any such j, $\vdash H(x+u+\widetilde{2},\widetilde{j})$ or $\vdash H(x+u+\widetilde{2},\widetilde{u})$ for some u satisfying jRu. By transitivity of R, we have mRu.

Thus in all cases, (A) or (B), we concluded that we have $\vdash H(x+u+\widetilde{2},\widetilde{m})$ or $\vdash H(x+u+\widetilde{2},\widetilde{k})$ for some k such that mRk.

□

COROLLARY C.5.16. $\vdash_{\mathcal{T}} H(x,\widetilde{m}) \to S_m \vee \bigvee_{mRj} S_j$.

The usual proof found in the literature takes advantage of the fact that the converse relation R^{-1} is well-founded hence if we have $aR^{-1}b$ then we have an R^{-1}-minimal a' such that $a'R^{-1}b$. So one can do induction along R^{-1}.

Of course, bRa' and a' is R-maximal. Now some of these proofs start by saying "we do an informal induction on the converse of R".

We will almost do the same but will not say "informal induction". R is finite and the proof that looks like induction is actually a "packing" in induction-like language a tedious induction-free proof that proceeds backwards along the graph (W, R). The proof *is* formal.

PROOF: Cases according to m:

1. m is maximal. Then no j satisfies mRj hence the $\bigvee_{mRj}$ part of the disjunction in (¶) (C.5.15) drops out.

   We simply have $\vdash_{\mathcal{T}} H(x, \widetilde{m}) \to (\forall z > x)H(z, \widetilde{m})$ which in shorthand is $\vdash_{\mathcal{T}} H(x, \widetilde{m}) \to S_m$.

2. m is not maximal. (¶) states

   $$\vdash_{\mathcal{T}} H(x, \widetilde{m}) \to (\forall z > x)\Big(H(z, \widetilde{m}) \vee \bigvee_{mRj} H(z, \widetilde{j})\Big)$$

   Now, for all such a j —which are "below m relative to R^{-1}"— we have already *proved*

   $$\vdash_{\mathcal{T}} H(x, \widetilde{j}) \to S_j \vee \bigvee_{jRk} S_j \qquad (D)$$

   walking backwards along the graph (we are *not* saying "by R induction hypothesis").

   By (D) and (¶) via tautological implication and $\forall$-*monotonicity*[48] we get

   $$\vdash_{\mathcal{T}} H(x, \widetilde{m}) \to (\forall z > x)\Big(H(z, \widetilde{m}) \vee \bigvee_{mRj} \Big(S_j \vee \bigvee_{jRk} S_k\Big)\Big)$$

   By transitivity of R the part "$\bigvee_{mRj} \Big(S_j \vee \bigvee_{jRk} S_k\Big)$" above simplifies to"$\bigvee_{mRk} S_k\Big)$", that is, we have,

   $$\vdash_{\mathcal{T}} H(x, \widetilde{m}) \to (\forall z > x)\Big(H(z, \widetilde{m}) \vee \bigvee_{mRk} S_k\Big)$$

[48] That is, $A \to B \vdash (\forall x)A \to (\forall x)B$; cf. [38].

$$\vdash_{\mathcal{T}} H(x,\widetilde{m}) \rightarrow \Big((\forall z > x) H(z,\widetilde{m}) \Big) \vee \bigvee_{mRk} S_k$$

$$\vdash_{\mathcal{T}} H(x,\widetilde{m}) \rightarrow S_m \vee \bigvee_{mRk} S_k \qquad \text{the big brackets above says "} S_m \text{"}$$

The second step above is due to the pure logic theorem "$(\forall x)(A \vee B) \leftrightarrow ((\forall x)A) \vee B$ if x is not free in B".

□

COROLLARY C.5.17. $\vdash_{\mathcal{T}} (\exists x) H(x,\widetilde{m}) \rightarrow S_m \vee \bigvee_{mRj} S_j$.

PROOF: Directly from C.5.16 Directly from C.5.16 we have

$$\vdash_{\mathcal{T}} (\exists x) H(x,\widetilde{m}) \rightarrow (\exists x) \Big(S_m \vee \bigvee_{mRj} S_j \Big)$$

by $\exists$-monotonicity.[49] But $\vdash (\exists x)(S_m \vee \bigvee_{mRj} S_j) \leftrightarrow S_m \vee \bigvee_{mRj} S_j$ in pure logic since $S_m \vee \bigvee_{mRj} S_j$ has no free variables. Thus, $\vdash_{\mathcal{T}}$ $(\exists x) H(x,\widetilde{m}) \rightarrow S_m \vee \bigvee_{mRj} S_j$. □

COROLLARY C.5.18. $\vdash_{\mathcal{T}} S_0 \vee S_1 \vee \ldots \vee S_n$.

PROOF: By C.5.16,

$$\vdash_{\mathcal{T}} H(0,0) \rightarrow S_0 \vee \bigvee_{0Rj} S_j$$

The right hand side of the "$\rightarrow$" can be written as $S_0 \vee S_1 \vee \ldots \vee S_n$ since 0 reaches all $j \neq 0$ in our Kripke frame $\langle W, R \rangle$. The contention follows by modus ponens since $\vdash_{\mathcal{T}} H(0,0)$. □

LEMMA C.5.19 (Solovay Lemma C.5.2, (3)). *For any two i, j such that iRj we have $\vdash_{\mathcal{T}} S_i \rightarrow \neg\Theta(\ulcorner \neg S_j \urcorner)$.*

Note. For the rest of the section we write $\Theta(x)$ —without subscript— for the provability predicate for $\mathcal{T}$, that is, $(\exists y) Proof(y,x)$.

PROOF: Assume S_i and prove $\neg\Theta(\ulcorner \neg S_j \urcorner)$, or better still, add also $\Theta(\ulcorner \neg S_j \urcorner)$ to the hypothesis and obtain a contradiction.

1. Hypothesis S_i abbreviates $(\exists u)(\forall x \geq u) H(x,\widetilde{i})$. So let c be a *new* constant (cf. A.0.3) and add the hypothesis

$$(\forall x \geq c) H(x,\widetilde{i}) \qquad (1)$$

[49] $A \rightarrow B \vdash (\exists A) \rightarrow (\exists x) B$ of pure logic ([38])

2. Hypothesis $\Theta(\ulcorner\neg S_j\urcorner)$ abbreviates $(\exists y)Proof(y,\ulcorner\neg S_j\urcorner)$. Since a proof can be lengthened by repeating its last formula (the "theorem" it proves) we have $(\exists y \geq c)Proof(y,\ulcorner\neg S_j\urcorner)$ using the c from 1. above.

   Let b be a *newer* constant, and add the hypothesis

$$b \geq c \wedge Proof(b,\ulcorner\neg S_j\urcorner) \tag{2}$$

Now (1) plus Ax2 (4.1.1) and MP give

$$b \geq c \to H(b,\widetilde{i})$$

From (2) above (second conjunct) and the definition of H ((1) on p.79; see also (2) on p.80), combining with the assumption iRj, we also get $H(b,\widetilde{j})$, contradicting C.5.10. □

LEMMA C.5.20. *For $i \geq 1$ we have $\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner\neg S_i\urcorner)$.*

PROOF: Add the hypothesis S_i to $\mathcal{T}$. That is, we add $(\exists u)(\forall x > u)H(x,\widetilde{i})$.

By the least principle (equivalent to induction) and the above we have (in $\mathcal{T}+S_i$):

$$(\exists u)\Big((\forall x \geq u)H(x,\widetilde{i}) \wedge (\forall w < u)(\exists x \geq w)\neg H(x,\widetilde{i})\Big)$$

Let c be a new constant and add the following to our existing hypotheses:

$$(\forall x \geq c)H(x,\widetilde{i}) \wedge (\forall w < c)(\exists x \geq w)\neg H(x,\widetilde{i}) \tag{1}$$

Note that by $i \geq 1$, our hypotheses prove $c > 0$ else the limit is S_0 by the first conjunct of (1) and the theorem $H(0,0)$. This contradicts the assumption S_i (C.5.14). Now (first conjunct of (1) and Ax2 of 4.1.1; using MP) we have

$$H(c,\widetilde{i}) \tag{2}$$

The second conjunct similarly yields via Ax2

$$(\exists x \geq c \dot{-} \widetilde{1})\neg H(x,\widetilde{i})$$

Only $x = c \dot{-} \widetilde{1}$ works above else we contradict (1) (first conjunct). Thus we proved

$$\neg H(c \dot{-} \widetilde{1},\widetilde{i}) \tag{3}$$

(2) and (3) and definition of $H(c, \widetilde{i})$ entail that we are in case #2 of the definition (p.79), thus we have proved $Proof(c, notlim(\ulcorner H(x,y) \urcorner, \widetilde{i}))$ or (cf. (2) on p.80)

$$Proof(c, \ulcorner \neg S_i \urcorner)$$

Thus, trivially, we have also proved $\Theta(\ulcorner \neg S_i \urcorner)$. □

LEMMA C.5.21.

1. *S_0 is true in the standard model of PA ($\mathcal{T}$).*
2. *$\mathcal{T} + S_i$ is consistent, for $i = 0, 1, \ldots, n$. This is C.5.2,(2).*

PROOF: Both 1. and 2. are argued in the *metatheory* of PA ($\mathcal{T}$).

1. Let $i \geq 1$. By
$$\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner \neg S_i \urcorner) \tag{1}$$
we conclude that S_i is false: Otherwise (1) implies —by soundness — that $\Theta(\ulcorner \neg S_i \urcorner)$ is true. But that says that $\neg S_i$ is provable and thus by soundness it is true, contradicting the assumption that it is S_i that is true.

   Now soundness and $\vdash_{\mathcal{T}} S_0 \vee S_1 \vee \ldots \vee S_n$ implies that $S_0 \vee S_1 \vee \ldots \vee S_n$ is true. But all the S_i with $i \geq 1$ are false. So S_0 is true.

   Interesting by-product: So the limit of H is actually 0.

2. $\mathcal{T} + S_0$ is consistent since S_0 and all axioms of $\mathcal{T}$ are true in the same (standard) model.

   Now $\vdash_{\mathcal{T}} S_0 \to \neg\Theta(\ulcorner \neg S_i \urcorner)$, $i \geq 1$, and since S_0 is true, so is $\neg\Theta(\ulcorner \neg S_i \urcorner)$. But that says that $\nvdash_{\mathcal{T}} \neg S_i$, that is, it says that $\mathcal{T} + S_i$ is consistent.

□

The final lemma requires us to use the derivability conditions DC 1 and DC 2 (p.46). It also requires the use of

$$\vdash_{\mathcal{T}} A \to \Theta(\ulcorner A \urcorner)$$

that holds if A is semi-recursive (Σ_1) —cf. [38] for a proof that is extremely long just as is the proof for DC 1, DC 2.

LEMMA C.5.22 (C.5.2,(4)). *If $i > 0$, then $\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner \bigvee_{iRj} S_j \urcorner)$.*

PROOF:

1. Applying DC 1 and then DC 2 to C.5.17 we get

$$\vdash_{\mathcal{T}} \Theta(\ulcorner(\exists x)H(x,\widetilde{m})\urcorner) \to \Theta(\ulcorner S_m \vee \bigvee_{mRj} S_j\urcorner)$$

2. $\vdash_{\mathcal{T}} S_i \to (\exists x)H(x,\widetilde{i})$ trivially, since S_i says $(\exists u)(\forall x \geq u)H(x,\widetilde{i})$.[50]

3. $\vdash_{\mathcal{T}} (\exists x)H(x,\widetilde{i}) \to \Theta(\ulcorner(\exists x)H(x,\widetilde{i})\urcorner)$ because $H(x,\widetilde{i})$ being primitive recursive, $(\exists x)H(x,\widetilde{i})$ is Σ_1.

4. We have $\vdash_{\mathcal{T}} \neg S_i \to S_i \vee \bigvee_{iRj} S_j \to \bigvee_{iRj} S_j$ by Boolean logic.

Now Result 4. (above) yields (DC 1 followed by DC 2)

$$\vdash_{\mathcal{T}} \Theta(\ulcorner\neg S_i\urcorner) \to \Theta(\ulcorner S_i \vee \bigvee_{iRj} S_j\urcorner) \to \Theta(\ulcorner\bigvee_{iRj} S_j\urcorner) \tag{5}$$

In a sequence of steps that follow we will eliminate the first two "Θ" from (5).

- C.5.20 and (5) via tautological implication yield

$$\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner S_i \vee \bigvee_{iRj} S_j\urcorner) \to \Theta(\ulcorner\bigvee_{iRj} S_j\urcorner) \tag{5'}$$

  This step used the hypothesis $i \geq 1$.

- (5′) plus Result 1. and tautological implication yield

$$\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner(\exists x)H(x,\widetilde{m})\urcorner) \to \Theta(\ulcorner\bigvee_{iRj} S_j\urcorner) \tag{5''}$$

- (5″) plus Result 3. and Boolean logic yield

$$\vdash_{\mathcal{T}} S_i \to (\exists x)H(x,\widetilde{m}) \to \Theta(\ulcorner\bigvee_{iRj} S_j\urcorner) \tag{5'''}$$

[50] Introduce a new constant c, then add the hypothesis $(\forall x \geq c)H(x,\widetilde{i})$. Thus we have $H(c,\widetilde{i})$ by Ax2 and hence also $(\exists x)H(x,\widetilde{i})$

- $(5''')$ plus Result 2. and tautological implication yield

$$\vdash_{\mathcal{T}} S_i \to S_i \to \Theta(\ulcorner \bigvee_{iRj} S_j \urcorner)$$

  Which by Boolean logic we simplify to the concluding statement of the lemma.

□

Bibliography

[1] *Kurt Gödel Collected Works Volume 1: Publications 1929-1936*. Oxford University Press, Inc., 1986.

[2] S. Artemov and G. Dzhaparidze. Finite Kripke Models and Predicate Logics of Provability. *Journal of Symbolic Logic*, 55(3):1090–1098, 1990.

[3] Arnon Avron. On modal systems having arithmetical interpretations. *Journal of Symbolic Logic*, 49(3):935–942, 1984.

[4] George Boolos. *The logic of provability*. Cambridge University Press, 2003.

[5] R. Carnap. *Logische Syntax der Sprache*. Springer-Verlag, 1934.

[6] M. Davis. *The Undecidable*. Raven Press, Hewlett, NY, 1965.

[7] Martin Davis. *Computability and Unsolvability*. McGraw-Hill, New York, 1958.

[8] R. Dedekind. *Was sind und was sollen die Zahlen?* Vieweg, Braunschweig, 1888. (In English translation by W.W. Beman [9]).

[9] R. Dedekind. *Essays on the Theory of Numbers*. Dover Publications, New York, 1963. (First English edition translated by W.W. Beman and published by Open Court Publishing, 1901).

[10] Dick de Jongh, Marc Jumelet and Franco Montagna. On the Proof of Solovay's Theorem. *Studia Logica*, 50(1):51–69, 1991.

[11] Edsger W. Dijkstra and Carel S. Scholten. *Predicate Calculus and Program Semantics*. Springer, New York, 1990.

[12] Feng Gao and George Tourlakis. A short and readable proof of cut elimination for two first-order modal logics. *Bulletin of the Section of Logic*, 44(3/4), 2015.

[13] James Garson. Modal Logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edition, 2021.

[14] K. Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Math. und Physik*, 38:173–198, 1931. (Also in English in Davis [6, 5–38]).

[15] Kurt Gödel. Eine interpretation des intuitionistischen aussagenkalkuls. *Ergebnisse Math*, 4:39–40, 1933.

[16] Robert Goldblatt, Dov M. Gabbay, and John Woods. Logic and the modalities in the twentieth century. *Handbook of the History of Logic*, page 8–16, 2006.

[17] David Gries and Fred B. Schneider. *A Logical Approach to Discrete Math.* Springer, New York, 1994.

[18] David Gries and Fred B. Schneider. Adding the Everywhere Operator to Propositional Logic. *Journal of Logic and Computation*, 8(1):119–129, 1998.

[19] A. Grzegorczyk. Some classes of recursive functions. *Rozprawy Matematyczne*, 4:1–45, 1953.

[20] D. Hilbert and W. Ackermann. *Principles of Mathematical Logic.* Chelsea, New York, 1950.

[21] D. Hilbert and P. Bernays. *Grundlagen der Mathematik I and II.* Springer, New York, 1968.

[22] George Edward. Hughes and Maxwell J. Cresswell. *A new introduction to modal logic.* Routledge, 2007.

[23] G. Japaridze and D. de Jongh. The Logic of Provability. In Buss, S. R., editor, *Handbook of Proof Theory*, pages 475–550. Elsevier Science B.V., 1998.

[24] F. Kibedi and G. Tourlakis. A modal extension of weak generalisation predicate logic. *Logic Journal of IGPL*, 14(4):591–621, 2006.

[25] S.C. Kleene. *Introduction to metamathematics.* North-Holland, Amsterdam, 1952.

[26] Stephen C. Kleene. *Introduction to metamathematics.* Wolters-Noordhoff, 1974.

[27] Saul A. Kripke. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24(1):1–14, 1959.

[28] Elliott Mendelson. *Introduction to Mathematical Logic.* Wadsworth & Brooks, Monterey, CA, 3rd edition, 1987.

[29] Franco Montagna. The predicate modal logic of provability. *Notre Dame Journal of Formal Logic*, 25:179–189, 1984.

[30] H. Rogers. *Theory of Recursive Functions and Effective Computability.* McGraw-Hill, New York, 1967.

[31] Yehuda Schwartz and George Tourlakis. On the proof-theory of two formalisations of modal first-order logic. *Studia Logica*, 96(3):349–373, 2010.

[32] Yehuda Schwartz and George Tourlakis. On the proof-theory of a first-order extension of GL. *Logic and Logical Philosophy*, 23(3):329–363, 2013.

[33] Yehuda Schwartz and George Tourlakis. A proof theoretic tool for first-order modal logic. *Bulletin of the Section of Logic*, 42(3/4):93–110, 2013.

[34] Joseph R. Shoenfield. *Mathematical Logic.* Addison-Wesley, Reading, MA, 1967.

[35] C Smorynski. *Self-Reference and Modal Logic.* Springer, New York, 1985.

[36] R. M. Smullyan. *Gödel's Incompleteness Theorems*. Oxford University Press, New York, 1992.

[37] Robert M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 25(3–4):287–304, 1976.

[38] G. Tourlakis. *Lectures in Logic and Set Theory, Volume 1: Mathematical Logic*. Cambridge University Press, Cambridge, 2003.

[39] G. Tourlakis. *Mathematical Logic*. John Wiley & Sons, Hoboken, NJ, 2008.

[40] G. Tourlakis. *Theory of Computation*. John Wiley & Sons, Hoboken, NJ, 2012.

[41] G. Tourlakis. A new arithmetically incomplete first-order extension of GL all theorems of which have cut free proofs. *Bulletin of the Section of Logic*, 45(1):17–31, 2016.

[42] George Tourlakis and Francisco Kibedi. A modal extension of first order classical logic. part i. *Bulletin of the Section of Logic*, 32(4):165–178, 2003.

[43] George Tourlakis and Francisco Kibedi. A modal extension of first order classical logic. part ii. *Bulletin of the Section of Logic*, 33:1–10, 2004.

[44] V. A. Vardanyan. Arithmetic complexity of predicate logics of provability and their fragments. *Soviet Mathematics Doklady*, 34:384–387, 1986.

[45] Rostislav E. Yavorsky. On arithmetical completeness of first-order logics of provability. *Advances in Modal Logic*, pages 1–16, 2002.