Bulletin of the **S**ection of **L**ogic
Volume 50/4 (2021), pp. 513–541
https://doi.org/10.18778/0138-0680.2021.18

Yunge Hao
George Tourlakis

AN ARITHMETICALLY COMPLETE PREDICATE MODAL LOGIC

Abstract

This paper investigates a first-order extension of GL called ML^3. We outline briefly the history that led to ML^3, its key properties and some of its toolbox: the *conservation theorem*, its cut-free Gentzenisation, the "formulators" tool. Its semantic completeness (with respect to finite reverse well-founded Kripke models) is fully stated in the current paper and the proof is retold here. Applying the Solovay technique to those models the present paper 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—"$\Box$" simulating the the informal classical "$\vdash$"—is also arithmetically complete in the Solovay sense.

Keywords: Predicate modal logic, arithmetical completeness, logic GL, Solovay's theorem, equational proofs.

1. Introduction

Solovay introduced in [23] the propositional provability logic GL (Gödel-Löb 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*.

Presented by: Luiz Carlos Pereira
Received: October 17, 2020
Published online: August 23, 2021

There has been a lot of interest in discovering first-order provability logics (cf. [3]). The obvious idea seemed to be defining *extensionally* a "Quantified GL"—or QGL_1[1]—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 formulae A.

$$QGL_1 = \{P : \vdash_{PA} f(P) \text{ for } \textit{every} \text{ arithmetical interpretation } f\} \quad (1.1)$$

Vardanyan however showed ([29]) that this first-order logic is not recursively axiomatisable; in fact he proved a stronger result: The QGL_1 of (1.1) is Π^0_2-complete.

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 (1.1), with the same behaviour of $\Box$ vis a vis free variables, as that of QGL_1.

This *intentional* logic QGL, 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. [2, 17, 26]). It turns out that this QGL has shortcomings too:

- It has no cut-free *Gentzenisation*, i.e., no cut-free Gentzen-equivalent logic (cf. [2]).

- 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 [17].

In [1, 12] a first-order QGL-like extension of GL is investigated and proved to be arithmetically complete. *However*, while on one hand $\Box$ was still "transparent" to free variables, on the other the "finite" Kripke models were overly restrictive: The *domains* of each world were required to be *finite* as well and to satisfy certain inclusion relations at that.

Later on, Yavorsky [30] 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. QGL^b is recursively axiomatised and

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

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.

A closely related first-order logic is the ML^3 of [19], with essentially the same set of axioms and also with an "opaque" $\Box$, but for technical reasons this has only the first two of the above rules as primary, "hiding" necessitation in the axioms in the style of [22], 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 provability predicate modal logics 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 [30] note the detrimental effects of a "transparent" $\Box$ on arithmetical completeness. Thus, while the earlier paper of [1] obtained arithmetical completeness of a predicate modal logic *with* a transparent $\Box$ it did so *on the condition that* such a logic had severely restricted finite Kripke models ("finite" applying to the *domains* of said models as well).

Yavorski [30] successfully experimented with an *opaque* $\Box$ and with the restricted Barkan formula

$$\Box A \to \Box(\forall x)A \tag{1.2}$$

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 (1.2) (see however [27, 28, 26] who chose this axiom for reasons totally unrelated to arithmetical comleteness).

Through a different route, with some interesting intermediate stops, [27, 28, 19] arrived at the logic ML^3 that is the focus of the present paper, while [26] further explored the significance of axiom (1.2) in ML^3 and M^3, in particular proving

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

Thus, all other axioms being left *as is*, (1.2) is *essential* for arithmetical completeness.

ML^3 has an interesting and consistent history. [27, 28] introduced M^3 in response to a problem stated in [9]. The authors of the latter noted that formal (classical) equational proofs

$$A_1 \Leftrightarrow {}^{2} A_2 \Leftrightarrow \ldots \Leftrightarrow A_i \quad \overbrace{\vdash\dashv}^{\text{metatheoretical step}} \quad (\forall x)A_i \Leftrightarrow \ldots \Leftrightarrow A_n$$

must be necessarily *disconnected* at the step above where we want to state "A_i iff $(\forall x)A_i$". The step is metatheoretical because the formal $A_i \Leftrightarrow (\forall x)A_i$ is invalid, in particular $A_i \to (\forall x)A_i$ is. Thus [9] asked: *Given that* $\vdash A_i$ *iff* $\vdash (\forall x)A_i$ *holds in the metatheory, can we recast the equational proof above* within modal logic *like this*

$$\Box A_1 \Leftrightarrow \Box A_2 \Leftrightarrow \ldots \Leftrightarrow \Box A_i \Leftrightarrow \Box(\forall x)A_i \Leftrightarrow \ldots \Leftrightarrow \Box A_n$$

where $\Box$ *means classical provability* ($\vdash$), *and thus make all classical equational proofs so translated both formal (within modal logic) and also* connected?

[27, 28] answered this question affirmatively, building the first-order modal logic M^3 and proving *semantically* (via Kripke models) their *conservation theorem* which, essentially, states

$$A \vdash B \text{ classically iff } \vdash \Box A \to \Box B \text{ modally} \tag{1.3}$$

M^3 is a first-order extension of the propositional modal logic K4, and was introduced to satisfy (1.3), that is, to be a "provability logic" for *pure classical predicate logic* rather than for PA. Such a provability logic is especially useful in the practice of *equational proofs* of [4, 8, 25].

[27, 28] and the related [13] contain several examples of disconnected classical equational proofs that (1.3) helps to convert into *connected modal translations* of the former proofs.

There were two key design criteria for M^3:

- $\Box$ 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. [16, 21, 24]) we have $A \vdash (\forall x)A$. In the words of [27, 28],

  > *The motivation regarding [free] object variables [in* $\Box A$*] is our intended intuitive interpretation of* $\Box$ *as the classical* $\vdash$*, and therefore as the classical* $\models$ *as well. When we say "*$\models A$*" classically, we mean that for all structures where*

[2] Conjunctional formal equivalence. That is, $A \Leftrightarrow B \Leftrightarrow C$ is defined to mean $A \equiv B$ *and* $B \equiv C$.

we interpret A, and for all value-assignments to the free object variables of A, the formula is true. Thus the variables in a statement such as "$\models A$" are implicitly universally quantified and are unavailable for substitutions.

- We have strong generalisation in M^3 (and ML^3), that is $A \vdash (\forall x)A$, and thus we *must* have, by (1.3), the special case $\Box A \to \Box(\forall x)A$—the (1.2) above—as a (modal) theorem. The easy approach to *have* this special case as a theorem was to adopt it as an axiom. It was not known to the authors of [27, 28, 13, 19] at the time whether (1.2) was independent of the remaining axioms. This was established to be the case by one of the authors later [26].

Thus the above (original) interpretation of $\Box$ in M^3 and its extension ML^3 is totally different from the interpretation of the $\Box$ in GL. The box operator of GL is interpreted *arithmetically* as, essentially, $\Theta(x)$, defined below in this paragraph. The interpretation mapping is usually denoted by *. Thus, by induction on the formation of GL formulae, atomic formulae A of GL are mapped to arbitrarily chosen sentences A^* of PA. For the induction step * *commutes* with $\neg$ and $\wedge$, that is, $(\neg A)^*$ is $\neg A^*$ and $(A \wedge B)^*$ is $A^* \wedge B^*$. Finally, $(\Box A)^*$ is interpreted as $\Theta(\ulcorner A^* \urcorner)$—which says "$A^*$" is a PA-theorem—where $\ulcorner X \urcorner$ denotes the Gödel number of X [22, 7]. The Σ_1-formula $\Theta(x)$ stands for $(\exists y)Pr(y, x)$ where $Pr(y, x)$ is true iff the Gödel number y codes a PA-proof of the formula with Gödel number x. Thus $\Theta(x)$ is true iff x is the Gödel number of a theorem of PA.

The logic ML^3 was introduced in [19], adding Löb's axiom $\Box(\Box A \to A) \to \Box A$ to M^3,[3] thus it is a first-order extension of both GL and M^3, and hence can (provably) simulate classical provability $\vdash$ through $\Box$ as well. ML^3 is over the same language as its predecessor M^3, and in particular, $\Box A$ is closed for all A.

[18, 19] developed the proof theory for M^3 and ML^3 by devising cut-free Gentzenisations of each, called GTKS and GLTS respectively. They gave completely detailed proofs of the admissibility of cut in each logic. 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 *subformula property* of cut-free Gentzen proofs.

[3] While Łöb's axiom can prove the axiom $\Box A \to \Box\Box A$ of M^3, we will retain it here for technical convenience as was done in [19].

In fact one of the results in the aforementioned references was a proof-theoretic (syntactic) proof of (1.3).

We also note in this historical review that [6] devised significantly shorter proofs than those in [18, 19] for the admissibility of cut in each of M^3 and ML^3.

[20] introduced certain formula to formula mappings named *formulators* (*formu*la trans*lators*). Such mappings preserve proofs in logics such as M^3, ML^3, and QGL, that is, if $\Gamma \vdash A$ holds in any one of these logics, then for any well-chosen formulator $\mathfrak{F}$ in each case we can have $\mathfrak{F}(\Gamma) \vdash \mathfrak{F}(A)$. 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 ([2]), the formulators tool was applied profitably ([20, 26]).

For completeness sake, here is the definition of a formulator mapping $\mathfrak{F}$:

DEFINITION 1.1 (Formulators [20, 26]). A *formula translator* or *formulator* is a mapping, $\mathfrak{F}$, from the set of formulae over a modal language L to itself such that:

1. $\mathfrak{F}(A) = A$ for every *atomic* formula A.
2. $\mathfrak{F}(A \to B) = \mathfrak{F}(A) \to \mathfrak{F}(B)$ for all formulae A, B.
3. $\mathfrak{F}((\forall x)A[x]) = (\forall x)B[x]$, where $B[a] = \mathfrak{F}(A[a])$.
4. The free variables of $\mathfrak{F}(\Box A)$ are among those of $\Box A$.

   Thus $\mathfrak{F}(\Box A)$ can be almost anything, subject to the restriction stated. $\Box$

[19] 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 [23] which heavily hinges on such finite Kripke models, the authors conjectured the arithmetical completeness of ML^3 in the conclusions section (cf. also the introduction section of [26]).

The present paper proves this conjecture, adapting the idea from [30] 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.

2. Language and symbols

We will not go over the well known inductive definition of formulae over a first order alphabet (cf. [21, 24, 27]),[4] but we will note our notational conventions.

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.

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. "*The*" is justified since we can reorder the quantification sequence and also eliminate repetitions without affecting either the meaning or the provability of the closure.

We call a formula $\Box A$ *boxed*. It is always a sentence (cf. [26] for the exact syntax of $\Box A$). 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 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* $B \Leftrightarrow C$".

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. $A(x, u, w)$ coveys that x, u, w are *all* the free variables of A while $A[x, u, w]$ conveys that x, u, w *may* be free in A. In the former case we may write $A(\mathbf{i}, u, w)$, in the latter $A[\mathbf{i}, u, w]$, to indicate the result of $A[x := \mathbf{i}]$.

[4] Note that as a technical convenience towards effecting Gentzenisation, [19] separates object variables into *free* and *bound* types. Here we follow the standard syntactic approach where bound vs. free is determined by how the variable is used syntactically.

3. The Logic ML^3

The language L of ML^3 in the present paper will have predicate symbols but no function symbols or constants. However, the language will later be augmented (cf. 4.2 and 4.5) to include *imported* constants.

DEFINITION 3.1 (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. 4.2 and 4.5 that refer to *imported constants*). The result $A[y]$ is undefined if "y is captured by a quantifier" as in, e.g., [24].

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 $\Box A \to \Box\Box A$.

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.

[19] has introduced and studied a variant of ML^3 above, *with* function and constant symbols and with equality (and its axioms) included. It is simpler—and customary ([1, 12, 30])—to discuss arithmetical completeness without these features. □

DEFINITION 3.2. The rules of inference of ML^3 are two, modus ponens (MP) $A, A \to B \vdash B$ and (strong) generalisation $A \vdash (\forall x)A$.[5]

[5] 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.

$\Gamma \vdash A$ (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. [27, 19, 26]).

Remark 3.3 (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, $\models_{taut} A_1 \to A_2 \to \ldots \to A_n \to B$. $\Box$

Axiom group A1 immediately implies

Theorem 3.4 (Proof by tautological implication). *If $A_1, A_2, \ldots, A_n \models_{taut} B$, then $A_1, A_2, \ldots, A_n \vdash_{ML^3} B$.*

For the following see [19, 26].

Theorem 3.5 (Weak Necessitation). *If $\Gamma \vdash_{ML^3} A$, where $\Gamma = \Gamma' \cup \Box\Gamma'$ or $\Gamma = \Box\Gamma'$, then $\Gamma \vdash_{ML^3} \Box A$.*

4. Kripke semantics

Kripke's possible worlds semantics [15] is the standard model theoretic approach to modal logic.

Definition 4.1 (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 $w'Rw''Rw''' \ldots$ $\Box$

Definition 4.2 (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, $(\forall w \in W)(w = w_0 \vee w_0 R w)$. $\Box$

Definition 4.3 (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: Every $k \in M_w$ is imported as a formal constant $\mathbf{k}$. The resulting language is denoted by $L(M_w)$ ([21, 24]).

Definition 4.4. If $A(x_1, \ldots, x_n)$ is over L, then $A(\mathbf{k_1}, \ldots, \mathbf{k_n})$ over $L(\mathbb{N})$ is a sentence with *parameters* from $\mathbb{N}$. □

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

Definition 4.5 (Extended Interpretation; forcing truth in a world.). Firstly, we interpret all the imported constants of $L(\mathbb{N})$:

$I_w(\mathbf{k}) = k \in \mathbb{N}$, for each $\mathbf{k} \in L(\mathbb{N})$.

Next, by induction on *closed* formulae of $L(\mathbb{N})$, for every $w \in W$:

1. $I_w(\phi(\mathbf{k_1}, \ldots, \mathbf{k_n})) = \mathbf{t}$ 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 $\mathbb{N}$.
2. $I_w(\neg A) = \mathbf{t}$ iff $I_w(A) = \mathbf{f}$ for any closed formula A of $L(\mathbb{N})$.
3. $I_w(A \wedge B) = \mathbf{t}$ iff $I_w(A) = \mathbf{t}$ and $I_w(B) = \mathbf{t}$, for any closed formulae A and B of $L(\mathbb{N})$.
4. $I_w((\forall x)A) = \mathbf{t}$ iff $I_w(A[x := \mathbf{k}]) = \mathbf{t}$ for all $k \in \mathbb{N}$, where $(\forall x)A$ is a sentence of $L(\mathbb{N})$.
5. $I_w(\Box A) = \mathbf{t}$ 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 4.6 (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. $\square$

DEFINITION 4.7 (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 X from assumptions Γ, in symbols $\Gamma \models X$, means that every model of Γ is also a model of X; metatheoretically we may indicate this definition by "$(\forall \mathcal{M})\Big(\models_{\mathcal{M}} \Gamma$ implies $\models_{\mathcal{M}} X\Big)$". $\square$

5. Semantic completeness

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

The *soundness* of ML^3 is proved in [19] and will be omitted. It states,

PROPOSITION 5.1. For any given set of modal formulae Γ and any modal formula A, $\Gamma \vdash A$ implies that $\Gamma \models A$, where semantics are over finite transitive and irreflexive Kripke structures.

The Consistency Theorem [21, 22, 24] 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

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

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 5.2 (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: ([28]) The proof in its entirety can be found in loc. cit. and we will not repeat it here. In outline, let $\mathcal{T}$ be a consistent closed modal *theory* over the language of M^3.[6] For example, if we take $\mathcal{T}$ to be (intentionally) ML^3, then $\mathcal{T} = \forall\Lambda \cup \Box\Lambda$.

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. [21, 24]). 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.

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

Lemma 5.3 (Main Semantic Lemma for M^3, [21, 24, 28]).
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 such completion. We *define* (cf. [22, 28]) a relation R on the set of all completions by

$$\Delta R \Delta' \text{ iff } \Delta\Box\,^{7} \subseteq \Delta'$$

This R is transitive ([22, 28, 19]). Thus we let $W = \{w_0\} \cup \{w_a : w_0 R w_a\}$, discarding all inaccessible completions. The next lemma (not proved here) is

For all modal sentences B over $L(\mathbb{N})$ we have $w_a \Vdash B$ iff *$B \in w_a$* $\quad(\dagger)$

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

[7] $\Delta\Box$ is defined in Section 2.

By (†) 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 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 (†). Thus $\mathcal{M}$ is a Kripke model of $\mathcal{T}$. □

We next prove in detail that

THEOREM 5.4. *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 [19] below (5.8 and 5.10) 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 Łöb's axiom. The details follow.

Remark 5.5. 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 due to $\vdash_{M^3} \Box\forall B \leftrightarrow \Box B$ and the equivalence theorem.[8] Indeed, in one direction, note $\vdash_{M^3} \Box\forall B \to \Box B$ using repeated use of axiom A2, followed by weak necessitation and then repeated application of A5. In the other direction note $\vdash_{M^3} \Box B \to \Box\forall B$ 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 (e.g., [12]).

DEFINITION 5.6 (Adequate set of formulae). 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 $\neg\neg X$. □

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

Definition 5.7. 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 *augmemted* 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.

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. $\square$

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 [19] we use the set $S(A)$ to help us "flag" the *finite* subset W^f of worlds W that *we intend to keep*. Thus we define:

Definition 5.8. 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 $w \cap S(A) = w' \cap S(A)$.[9] 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 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: $W^f \overset{\text{reset}}{=} \left\{\beta : \alpha_0 \widehat{R} \beta\right\}$.

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

Proposition 5.9. $\widehat{R}$ 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.

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

- *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.

- *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. $\Box$

Lemma 5.10. *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 from [19] and is provided here for easy access. Induction on the complexity of X. As in loc. cit. 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 5.8 (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, $\alpha_i \Vdash_S B[x := \mathbf{k}]$ for all $k \in \mathbb{N}$. By the I.H. all the $B[x := \mathbf{k}]$ 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 $\neg B[x := \mathbf{m}]$ 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 $\alpha_i \vdash B[x := \mathbf{k}]$, for *all* $k \in \mathbb{N}$. By deductive closure $B[x := \mathbf{k}] \in \alpha_i$—and by the I.H. $\alpha_i \Vdash_S B[x := \mathbf{k}]$—for *all* $k \in \mathbb{N}$. By 4.5, case 5, $\alpha_i \Vdash_S (\forall x)B$.

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

   - Suppose $\Box B \in \alpha_i$. Thus, using "$\Rightarrow$" conjunctionally (metatheoretically)

$$\begin{aligned} \Box B \in \alpha_i &\overset{5.8(1)}{\Rightarrow} (\forall \alpha_j)(\alpha_i \widehat{R} \alpha_j \to \forall B \in \alpha_j) \\ &\overset{I.H.}{\Rightarrow} (\forall \alpha_j)(\alpha_i \widehat{R} \alpha_j \to \alpha_j \Vdash_S \forall B) \\ &\overset{4.5\ 6.}{\Rightarrow} \alpha_i \Vdash_S \Box B \end{aligned}$$

- For the converse we proceed contrapositively.

  So let

$$\Box B \notin \alpha_i \tag{5.1}$$

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

$$T = \{\Box D_1, \forall D_1, \ldots, \Box D_m, \forall D_m, \Box B, \neg\forall B\} \tag{5.2}$$

for some m. We claim that T *is consistent*. Proceeding by contradiction, *suppose otherwise*. Then (proof by contradiction, followed by the deduction theorem) $\Box D_1, \forall D_1, \ldots, \Box D_m, \forall D_m \vdash_{ML^3} \Box B \to \forall B$. Thus $\Box D_1, \forall D_1, \ldots, \Box D_m, \forall D_m \vdash_{ML^3} \Box B \to B$ (from $\forall B \to B$) hence $\Box D_1, \ldots, \Box D_m, \forall D_m \vdash_{ML^3} \Box(\Box B \to B)$ by weak necessitation. Now by tautological implication (via Löb's axiom) we get $\Box D_1, \forall D_1, \ldots, \Box D_m, \forall D_m \vdash_{ML^3} \Box B$, which implies $\Box B \in \alpha_i$ since α_i is deductively closed and contains the premises. We have just contradicted the main hypothesis of this bullet.

$$\text{Let then } \alpha_j \text{ be a } \textit{completion} \text{ of the consistent } T \tag{5.3}$$

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

$$\alpha_j \nVdash_S \forall B \tag{5.4}$$

If we can argue that we have

$$\alpha_i \widehat{R} \alpha_j \tag{5.5}$$

then we are done since (5.4) and (5.5) 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 (5.5). For the other half we have (5.1) and also need that $\Box B \in \alpha_j$. This is true by (5.2) and (5.3). $\Box$

THEOREM 5.11 ([19]). *ML^3 is complete with respect to* finite *reverse well-founded Kripke models (irreflexive and transitive).*

PROOF: To summarise, start at (¶). 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 5.10 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 5.10 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 $A \notin \alpha_0$. □

6. Arithmetical completeness

The main tool in this section is Solovay's work [23]. We build on [19] but also use two tools from [30], namely, a definition and a lemma in loc. cit., which appear modified below as 6.7 and 6.8 respectively. Our induction in the proof of 6.9 proceeds in its details differently.

THEOREM 6.1 (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 in [23] (for GL) we prove 6.1 contrapositively: Thus, assume $\mathrm{ML}^3 \nvdash A$, for some fixed modal *sentence* A over L, 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

$$\alpha_0 \nVdash_S A \tag{§}$$

This was done in 5.11 above.

The second phase is to apply Solovay's technique [23] to embed $\mathcal{M}$ in an appropriate $\mathcal{T}$—*which is a* finite *extension of PA that we define below*—and propose an arithmetical *realisation* * such that $\mathcal{T} \nvdash A^*$.

An *a priori* requirement of the embedding is that the worlds α_i (cf. 5.8) 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[10] 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 $\widehat{R}^0 = \widehat{R} \cup \{(0, i) : 1 \le i \le n\}$
- 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 $1 \le i \le n$.
- $W^{f,0} = \{0, 1, 2, 3, \ldots, n\}$

It is this $\mathcal{M}^0$ that Solovay embeds into PA (or extension $\mathcal{T}$). Below we list the Solovay lemmata that, interestingly, *can be used here as is* without reference to their complex proofs (not so in [1, 12]). For simplicity of *use* and exposition, many authors ([3, 30, 1]) use the abbreviations S_k or σ_k for the formal sentence (in PA) "$\mathbf{l} = \widetilde{k}$" that is pervasive in [23], 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 $W^{f,0}$.

Lemma 6.2 (Solovay's Lemmata). *$\mathcal{T}$ is some recursive extension of PA over a finite extension of the PA language. There are sentences S_i, for $0 \le i \le n$, of the language, such that*

(1) *For all $i \ne j$, $\vdash_{\mathcal{T}} \neg S_i \vee \neg S_j$.*
(2) *For $0 \le i \le n$, $\mathcal{T} + S_i$ is consistent.*
(3) *If $i\widehat{R}^0 j$, then $\vdash_{\mathcal{T}} S_i \to \neg\Theta_{\mathcal{T}}(\ulcorner \neg S_j \urcorner)$, where $\Theta_{\mathcal{T}}$ is the provability predicate for $\mathcal{T}$.*
Under the given assumptions, [23] formulated this as the equivalent $\vdash_{\mathcal{T}} S_i \to Cons_{\mathcal{T}+S_j}$. 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 $\vdash_{PA} S_i \to \Theta_{\mathcal{T}}(\ulcorner \bigvee_{i\widehat{R}^0 j} S_j \urcorner)$.*

As in [30] we will work with a specific finite consistent extension $\mathcal{T}$ of PA rather than PA. Towards obtaining this theory, we build consistent sets of *classical* formulae $\mathcal{C}_i$ (6.4 below) as follows.

[10] The technical reason is simply that Solovay's Kripke-frame-walking function h must be *total*—in fact, with some care ([12]) h can be proved to be primitive recursive—indeed must be initialised as $h(0) = 0$. We do not use Solovay's S_0 in our proof, nor do we mention S_0 in Lemma 6.2. Incidentally, S_0 is true in the standard model of PA, but not provable in PA. Solovay and [3] use the truth of S_0 in proving arithmetical completeness of GL. [30] and [22] do not. We follow the latter's paradigm here.

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, then so is $X(y)$ and vice versa by virtue of $\vdash (\forall y)X \to X(y)$ absolutely (axiom A2) and $X(y) \vdash (\forall y)X$. Thus we depart from the worlds i of [19], only using finite parts of them to define (in 6.4 via 6.3) the *finite* classical sets $\mathcal{C}_i$. These sets are needed for Proposition 6.6 that leads to the *finite* extension of PA.

Definition 6.3. For each $1 \le i \le n$, $S^i_{\max}(A)$ denotes a *maximal consistent subset of* $S(A)$ that contains $i \cap S(A)$ ($= \alpha_{i-1} \cap S(A)$).[11] □

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

(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 $\bot$.[12] The transformed formula X—again given the name $X^{t,i}$—is placed in $\mathcal{C}_i$. □

Remark 6.5. "t" is for "transformed" formula. But why the extra superscript i? Because the same X may appear in i and j, for $i \ne j$. 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 6.6. $\mathcal{C}_i$ is consistent iff $S^i_{\max}(A)$ is consistent.

Proof: Let $X \in S^i_{\max}(A)$. Note that, if $\Box B \in i$, then $i \vdash \Box B \equiv \top$[13] while if $\Box B \notin i$, then $\neg\Box B$ is in i by maximality, hence $i \vdash \Box B \equiv \bot$.[14]

---

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

[12] Case of $\neg\Box X$ 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 $\bot$.

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

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

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 thus in $S^i_{\max}(A)$ and 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 $\bot$. $\Box$

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 [11, Vol. II] and [10, 14]) 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, 5.2) that occurs in every such X into the numeral $\widetilde{k}$ to obtain a formula $re_i(X)$ in the language of PA. 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 $\mathrm{PA} + re_i(\mathcal{C}_i)$ 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

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

Now the arithmetical realisation $*$ of modal formulae, as usual, maps all the subformulae X of A into formulae of PA in the standard manner, that is, $*$ commutes with the Boolean connectives and $(\forall x)$, it preserves the free variables of X, and also commutes with substitution of variables for variables, that is if $X(x_1, \ldots, x_m)^* = Y(x_1, \ldots, x_m)$, then $X(y_1, \ldots, y_m)^* = Y(y_1, \ldots, y_m)$. Lastly, $(\Box A)^* = \Theta(\ulcorner A^* \urcorner)$, where here and for the rest of the proof we write just "Θ" for "$\Theta_{\mathcal{T}}$".

Definition 6.7 (Arithmetical realisation; initialisation).

Let B be any *atomic* subformula of A, where A was fixed at the outset of this section (cf. (§)). Being atomic it is classical.

Then for the *basis* of the realisation $*$ we set ([30]),[15]

$$B^* \overset{Def}{\leftrightarrow} \bigvee_{\substack{1 \leq j \leq n \\ j \Vdash \forall B}} S_j \wedge re_j(B^{t,j}) \tag{6.1}$$

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

The following useful lemma is stated in Yavorsky [30] without proof. A proof is the following.

LEMMA 6.8. $\vdash_{\mathcal{T}} S_i \to (B^* \leftrightarrow re_i(B^{t,i}))$ *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

$$\text{we need to prove instead } \vdash_{\mathcal{T}+S_i} B^* \leftrightarrow re_i(B^{t,i}) \tag{6.2}$$

We now proceed with our induction on classical formulae B:

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

$$\vdash_{\mathcal{T}+S_i} B^* \leftrightarrow re_i(B^{t,i}) \vee \bigvee_{\substack{j \neq i \\ j \Vdash \forall B}} S_j \wedge re_j(B^{t,j}) \tag{6.3}$$

   Note that by 6.2(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 $\perp$). 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.

2. 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 $(\forall x)D^*$

[15] Recall the renaming of α_j as $j+1$, at the beginning of Section 6.

by the definition of * while, by the definition of re_i, $re_i\big((\forall x)D^{t,i}\big)$ is $(\forall x)re_i(D^{t,i})$. $\square$

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. [24] 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 6.9 (Main Lemma). *Having got a finite Kripke model of n-nodes such that $1 \nVdash_{S^0} A$ (cf. §), 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 $1 \le i \le n$, 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 6.8) $\vdash_{\mathcal{T}+S_i} X^* \leftrightarrow re_i(X^{t,i})$. 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 by (6.1) the disjunct $S_i \wedge re_{i-1}(X^{t,i-1})$ is missing. By item 1. in the proof of 6.8 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.[16] and definition by cases,

[16] We remind the reader that as in [19] $\Box B$ is more complex than $\forall B$.

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

Applying DC1 then DC2 followed by modus ponens,

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

By 6.2(4) $\vdash_{\mathcal{T}} S_i \to \Theta(\ulcorner \bigvee_{i\widehat{R}^0 j} S_j \urcorner)$ and hence, by ($*$),

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

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

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

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

By 6.2(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 (§§) derives

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

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, not proved in this paper), $\vdash_{\mathcal{T}} (\Box B)^* \to (\Box\forall B)^*$, that is, $\vdash_{\mathcal{T}} \Theta(\ulcorner B^* \urcorner) \to \Theta(\ulcorner (\forall B)^* \urcorner)$. From ($***$) and the preceding we now get $\vdash_{\mathcal{T}} S_i \to \neg\Theta(\ulcorner B^* \urcorner)$, 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 6.8 instead.

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

$$\vdash_{\mathcal{T}+S_i} (\forall x)B^* \leftrightarrow re_i((\forall x)B) \tag{6.4}$$

(a) Now, if $i \Vdash_{S^0} (\forall x)B$, then $\vdash_{\mathcal{T}} re_i((\forall x)B)$. Tautological implication and (6.4) 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., $\vdash_{\mathcal{T}} \neg re_i((\forall x)B)$. (6.4) 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

  $$\vdash_{\mathcal{T}+S_i} \big((\forall x)B'\big)^* \tag{6.5}$$

  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 (6.5).
- 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 $\perp$ on the ML^3 side and by $\perp^*$ 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

$$\vdash_{\mathcal{T}+S_i} \neg\big((\forall x)B'\big)^* \tag{6.6}$$

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 (6.6) 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. $\Box$

Proof of the main theorem. By 6.9, 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 6.2(2) $\mathcal{T}+S_1$ is consistent, hence so is $\mathcal{T}+\neg A^*$.[17] Thus $\nvdash_{\mathcal{T}} A^*$. $\Box$

7. Concluding note

As remarked in [19] and more recently in [26], ML^3, being a first-order extension of GL due to the inclusion of the Löb axiom (A6), was meant to be a possible *candidate* for a modal *first-order* provability logic for (arithmetised provability in) PA.

Secondly, it was deliberately built as an extension of M^3 in order to *remain* a provability logic for *classical pure first-order logic.*

Indeed, the *conservation theorem* was proved (syntactically) for ML^3 (as it was for M^3) in [19] verifying that the second design criterion was met.

Given the establishment of its *semantic completeness* with respect to reverse well-founded *finite* and transitive Kripke structures ([19], and also in this paper), [19, 26] conjectured that the first design criterion ought to be also met. A proof of this has been offered in the present paper.

This paper benefits from the idea in [30] to show arithmetical completeness with respect to a *finite extension of PA* and also from Lemma 6.8 which is only stated in [30] but it is proved here.

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

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 metatheoretical work be done directly, without Gentzenisation, using formulators to investigate the Hilbert-style axiom system 3.1—[20, 26]. The second of the preceding references shows that in the presence of all the other axioms, the addition of A7 is *essential* for arithmetical completeness, since *all* its arithmetical interpretations *are* provable in PA, but A7 is independent of the other axioms of ML^3 (and M^3).

Moreover, *Craig's Interpolation* holds both for the Gentzenisation GLTS of ML^3 and the GTKS of M^3 ([19]), a property that fails for predicate modal logics in general ([5]).

[30] does not remark on whether QGL^b admits a Gentzenisation (cut-free or otherwise) but more remarkably it does not discuss the central importance of A7 as an axiom towards arithmetical completeness.

The origins of QGL^b and ML^3 are quite distinct, as the former was built 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 was carefully explained in [27, 28]—see also the quotation from [27, 28] in the present paper, on p. 4, first bullet—where we also explicate the choice of A7 (second bullet) as the modal counterpart of the classical $A \vdash (\forall x)A$. A7 appears to have been adopted without any obvious rationale in [30], mentioned only in passing as an assumption on which the normal form of modal formulae is based (*loc. cit.*, remark below Definition 2.1 on p. 3).

References

[1] S. Artemov, G. Dzhaparidze, *Finite Kripke Models and Predicate Logics of Provability*, **Journal of Symbolic Logic**, vol. 55(3) (1990), pp. 1090–1098, DOI: https://doi.org/10.2307/2274475.

[2] A. Avron, *On modal systems having arithmetical interpretations*, **Journal of Symbolic Logic**, vol. 49(3) (1984), pp. 935–942, DOI: https://doi.org/10.2307/2274147.

[3] G. Boolos, **The logic of provability**, Cambridge University Press (2003), DOI: https://doi.org/10.1017/CBO9780511625183.

[4] E. W. Dijkstra, C. S. Scholten, **Predicate Calculus and Program Semantics**, Springer, New York (1990), DOI: https://doi.org/10.1007/978-1-4612-3228-5.

[5] K. Fine, *Failures of the interpolation lemma in quantfied modal logic*, **Journal of Symbolic Logic**, vol. 44(2) (1979), pp. 201–206, DOI: https://doi.org/10.2307/2273727.

[6] F. Gao, G. Tourlakis, *A Short and Readable Proof of Cut Elimination for Two First-Order Modal Logics*, **Bulletin of the Section of Logic**, vol. 44(3/4) (2015), DOI: https://doi.org/10.18778/0138-0680.44.3.4.03.

[7] K. Gödel, *Eine Interpretation des intuitionistischen Aussagenkalkuls*, **Ergebnisse Math**, vol. 4 (1933), pp. 39–40.

[8] D. Gries, F. B. Schneider, **A Logical Approach to Discrete Math**, Springer, New York (1994), DOI: https://doi.org/10.1007/978-1-4757-3837-7.

[9] D. Gries, F. B. Schneider, *Adding the Everywhere Operator to Propositional Logic*, **Journal of Logic and Computation**, vol. 8(1) (1998), pp. 119–129, DOI: https://doi.org/10.1093/logcom/8.1.119.

[10] D. Hilbert, W. Ackermann, **Principles of Mathematical Logic**, Chelsea, New York (1950).

[11] D. Hilbert, P. Bernays, **Grundlagen der Mathematik I and II**, Springer, New York (1968), DOI: https://doi.org/10.1007/978-3-642-86894-8.

[12] G. Japaridze, D. de Jongh, *The Logic of Provability*, [in:] S. R. Buss (ed.), **Handbook of Proof Theory**, Elsevier Science B.V. (1998), pp. 475–550, DOI: https://doi.org/10.1016/S0049-237X(98)80022-0.

[13] F. Kibedi, G. Tourlakis, *A Modal Extension of Weak Generalisation Predicate Logic*, **Logic Journal of IGPL**, vol. 14(4) (2006), pp. 591–621, DOI: https://doi.org/10.1093/jigpal/jzl025.

[14] S. Kleene, **Introduction to metamathematics**, North-Holland, Amsterdam (1952).

[15] S. A. Kripke, *A completeness theorem in modal logic*, **Journal of Symbolic Logic**, vol. 24(1) (1959), pp. 1–14, DOI: https://doi.org/10.2307/2964568.

[16] E. Mendelson, **Introduction to Mathematical Logic**, 3rd ed., Wadsworth & Brooks, Monterey, CA (1987), DOI: https://doi.org/10.1007/978-1-4615-7288-6.

[17] F. Montagna, *The predicate modal logic of provability*, **Notre Dame Journal of Formal Logic**, vol. 25(2) (1984), pp. 179–189, DOI: https://doi.org/10.1305/ndjfl/1093870577.

[18] Y. Schwartz, G. Tourlakis, *On the Proof-Theory of two Formalisations of Modal First-Order Logic*, **Studia Logica**, vol. 96(3) (2010), pp. 349–373, DOI: https://doi.org/10.1007/s11225-010-9294-y.

[19] Y. Schwartz, G. Tourlakis, *On the proof-theory of a first-order extension of GL*, **Logic and Logical Philosophy**, vol. 23(3) (2013), pp. 329–363, DOI: https://doi.org/10.12775/llp.2013.030.

[20] Y. Schwartz, G. Tourlakis, *A proof theoretic tool for first-order modal logic*, **Bulletin of the Section of Logic**, vol. 42(3/4) (2013), pp. 93–110.

[21] J. R. Shoenfield, **Mathematical Logic**, Addison-Wesley, Reading, MA (1967).

[22] C. Smorynski, **Self-Reference and Modal Logic**, Springer, New York (1985), DOI: https://doi.org/10.1007/978-1-4613-8601-8.

[23] R. M. Solovay, *Provability interpretations of modal logic*, **Israel Journal of Mathematics**, vol. 25(3–4) (1976), pp. 287–304, DOI: https://doi.org/10.1007/bf02757006.

[24] G. Tourlakis, **Lectures in Logic and Set Theory, Volume 1: Mathematical Logic**, Cambridge University Press, Cambridge (2003), DOI: https://doi.org/10.1017/CBO9780511615559.

[25] G. Tourlakis, **Mathematical Logic**, John Wiley & Sons, Hoboken, NJ (2008), DOI: https://doi.org/10.1002/9781118032435.

[26] 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**, vol. 45(1) (2016), pp. 17–31, DOI: https://doi.org/10.18778/0138-0680.45.1.02.

[27] G. Tourlakis, F. Kibedi, *A modal extension of first order classical logic. Part I*, **Bulletin of the Section of Logic**, vol. 32(4) (2003), pp. 165–178.

[28] G. Tourlakis, F. Kibedi, *A modal extension of first order classical logic. Part II*, **Bulletin of the Section of Logic**, vol. 33 (2004), pp. 1–10.

[29] V. A. Vardanyan, *Arithmetic complexity of predicate logics of provability and their fragments*, **Soviet Mathematics Doklady**, vol. 34 (1986), pp. 384–387, URL: http://mi.mathnet.ru/eng/dan8607.

[30] R. E. Yavorsky, *On Arithmetical Completeness of First-Order Logics of Provability*, **Advances in Modal Logic**, (2002), pp. 1–16, DOI: https://doi.org/10.1142/9789812776471_0001.

Yunge Hao
York University
Department of Electrical Engineering and Computer Science
M3J1P3, Keele St #4700
Toronto, Ontario, Canada
e-mail: hyggs@my.yorku.ca

George Tourlakis
York University
Department of Electrical Engineering and Computer Science
M3J1P3, Keele St #4700
Toronto, Ontario, Canada
e-mail: gt@cse.yorku.ca