Annals of Pure and Applied Logic 61 (1993) 75–93
North-Holland

On the provability logic of bounded arithmetic

Alessandro Berarducci*
Dipartimento di Matematica, Universitá di Pisa, Via Buonarrotti 2, 56100 Pisa, Italy

Rineke Verbrugge**
Faculteit der Wiskunde en Informatica, Universiteit van Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, Netherlands

Communicated by D. van Dalen
Received 27 September 1991
Revised 6 March 1992

Abstract

Berarducci, A. and R. Verbrugge, On the provability logic of bounded arithmetic, Annals of Pure and Applied Logic 61 (1993) 75–93.

Let $PL\Omega$ be the provability logic of $I\Delta_0 + \Omega_1$. We prove some containments of the form $L \subseteq PL\Omega \subset Th(\mathscr{C})$ where L is the provability logic of PA and $\mathscr{C}$ is a suitable class of Kripke frames.

1. Introduction

In this paper we develop techniques to build various sets of highly undecidable sentences in $I\Delta_0 + \Omega_1$. Our results stem from an attempt to prove that the modal logic of provability in $I\Delta_0 + \Omega_1$, here called $PL\Omega$, is the same as the modal logic L of provability in PA. It is already known that $L \subseteq PL\Omega$. We prove here some strict containments of the form $PL\Omega \subset Th(\mathscr{C})$ where $\mathscr{C}$ is a class of Kripke frames.

Stated informally the problem is whether the provability predicates of $I\Delta_0 + \Omega_1$ and PA share the same modal properties. It turns out that while $I\Delta_0 + \Omega_1$ certainly satisfies all the properties needed to carry out the proof of Gödel's second incompleteness theorem (namely $L \subseteq PL\Omega$), the question whether $L = PL\Omega$ might depend on difficult issues of computational complexity. In fact if

Correspondence to: R. Verbrugge, Faculteit der Wiskunde en Informatica, Universiteit van Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, Netherlands.
* Research partially supported by the Italian Research Projects 40% and 60%.
** Research supported by the Netherlands Organization for Scientific Research (NWO).

$PL\Omega \neq L$, it would follow that $I\Delta_0 + \Omega_1$ does not prove its completeness with respect to Σ_1^0-formulas, and a fortiori $I\Delta_0 + \Omega_1$ does not prove the Matijasevič–Robinson–Davis–Putnam theorem (every r.e. set is diophantine, see [6], [3]). On the other hand if $I\Delta_0 + \Omega_1$ did prove its completeness with respect to Σ_1^0-formulas, it would follow not only that $L = PL\Omega$, but also that $NP = co\text{-}NP$. The possibility remains that $L = PL\Omega$ and that one could give a proof of this fact without making use of provable Σ_1^0-completeness in its full generality. Such a project is not without challenge due to the ubiquity of Σ_1^0-completeness in the whole area of provability logic.

We begin by giving the definitions of L and $PL\Omega$.

Definition 1.1. The language of modal logic contains a countable set of propositional variables, a propositional constant $\perp$, boolean connectives $\neg$, $\wedge$, $\rightarrow$, and the unary modality $\Box$. The modal provability logic L is axiomatized by all formulas having the form of propositional tautologies (including those containing the $\Box$-operator) plus the following axiom schemes:

1. $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$.
2. $\Box(\Box A \rightarrow A) \rightarrow \Box A$.
3. $\Box A \rightarrow \Box\Box A$.

The rules of inference are:

1. If $\vdash A \rightarrow B$ and $\vdash A$, then $\vdash B$ (modus ponens).
2. If $\vdash A$, then $\vdash \Box A$ (necessitation).

Definition 1.2. Let T be a Σ_1^b-axiomatized theory in the language of arithmetic (see [1]). A *T-interpretation* $*$ is a function which assigns to each modal formula A a sentence A^* in the language of T, and which satisfies the following requirements:

1. $\perp^*$ is the sentence $0 = 1$.
2. $*$ commutes with the propositional connectives, i.e., $(A \rightarrow B)^* = A^* \rightarrow B^*$, etc.
3. $(\Box A)^* = Prov_T(\ulcorner A^* \urcorner)$.

Clearly $*$ is uniquely determined by its restriction to the propositional variables. The presence in the modal language of the propositional constant $\perp$ allows us to consider closed modal formulas, i.e., modal formulas containing no propositional variables. If A is closed, then A^* does not depend on $*$, e.g. $(\Box\perp)^*$ is the arithmetical sentence $Prov_T(\ulcorner 0 = 1 \urcorner)$.

Definition 1.3. Let $PL\Omega$ be the provability logic of $I\Delta_0 + \Omega_1$, i.e., $PL\Omega$ is the set of all those modal formulas A such that for all $I\Delta_0 + \Omega_1$-interpretations $*$, $I\Delta_0 + \Omega_1 \vdash A^*$.

It is easy to see that $PL\Omega$ is deductively closed (with respect to modus ponens and necessitation), so we can write $PL\Omega \vdash A$ for $A \in PL\Omega$. Our results arise from an attempt to answer the following:

Question 1.4. Is $PL\Omega = L$? (Where we have identified L with the set of its theorems.)

The soundness side of the question, namely $L \subseteq PL\Omega$, has already been answered positively. This depends on the fact that any reasonable theory which is at least as strong as Buss' theory S^1_2 satisfies the derivability conditions needed to prove Gödel's incompleteness theorems (provided one uses efficient coding techniques and employs binary numerals). For the completeness side of the question, namely $PL\Omega \subseteq L$, we will investigate whether we can adapt Solovay's proof that L is the provability logic of PA.

We assume that the reader is familiar with the Kripke semantics for L and with the method of Solovay's proof as described in [9]. In particular we need the following:

Theorem 1.5. *$L \vdash A$ iff A is forced at the root of every finite tree-like Kripke model. (It is easy to see that A will then be forced at every node of every finite tree-like Kripke model.)*

Solovay's method is the following: if $L \nvdash A$, then the countermodel $(K, <, \Vdash)$ provided by the above theorem is used to construct a PA-interpretation $*$ for which $\mathrm{PA} \nvdash A^*$.

The reason Solovay's proof cannot be adapted to $I\Delta_0 + \Omega_1$ is that it is not known whether $I\Delta_0 + \Omega_1$ satisfies provable Σ^0_1-completeness (see Definition 2.1) which is used in an essential way in Solovay's proof.

2. Arithmetical preliminaries

Definition 2.1. Let Γ be a set of formulas. We say that a (Σ^b_1-axiomatized) theory T satisfies *provable Γ-completeness,* if for every formula $\sigma(\boldsymbol{x}) \in \Gamma$,

$$T \vdash \sigma(x_1, \ldots, x_n) \rightarrow Prov_T(\ulcorner \sigma(\dot{x}_1, \ldots, \dot{x}_n) \urcorner).$$

It is known that PA, as well as any reasonable theory extending $I\Delta_0 + exp$, satisfies provable Σ^0_1-completeness.

De Jongh, Jumelet and Montagna [5] showed that Solovay's result can be extended to all reasonable Σ^0_1-sound theories T satisfying provable Σ^0_1-completeness. More precisely it is sufficient that the provability predicate of T provably satisfies the axioms of Guaspari and Solovay's modal witness comparison logic R^-. So Solovay's result holds for ZF, $I\Sigma_n$ and $I\Delta_0 + exp$.

On the other hand it is known that if $NP \neq co\text{-}NP$, then $I\Delta_0 + \Omega_1$ does not satisfy provable Σ_1^0-completeness or even provable Δ_0-completeness. In [13] the second author proved that, if $NP \neq co\text{-}NP$, $I\Delta_0 + \Omega_1$ does not even satisfy provable completeness for the single Σ_1^0-formula

$$\sigma(u, v) \equiv \exists x\, (\mathrm{Prf}_{I\Delta_0+\Omega_1}(x, u) \wedge \forall y < x\, \neg \mathrm{Prf}_{I\Delta_0+\Omega_1}(y, v)).$$

One possibility, although unlikely, remains: to adapt Solovay's proof to $I\Delta_0 + \Omega_1$ it would suffice that $I\Delta_0 + \Omega_1$ satisfies provable Σ_1^0-completeness for *sentences*, and we cannot rule out this possibility even assuming $NP \neq co\text{-}NP$. By [5] it would actually suffice to have provable Σ_1^0-completeness for all closed instances of $\sigma(u, v)$ where u and v are instantiated by Gödel numbers of arithmetical sentences.

In view of the above difficulties, we try to do without Σ_1^0-completeness. In the rest of this section we state some results about $I\Delta_0 + \Omega_1$ which in some cases allow us to dispense with the use of Σ_1^0-completeness. The following proposition is proved in [15]:

Theorem 2.2. *$I\Delta_0 + \Omega_1$ satisfies provable Σ_1^b-completeness.*

By abuse of notation we will denote by $\Box A$ both the arithmetization of the provability predicate of $I\Delta_0 + \Omega_1$ and the corresponding modal operator. $\Diamond A$ is defined as $\neg\Box\neg A$ and $\Box^+ A$ as $\Box A \wedge A$. If $A(x)$ is an arithmetical formula, we will write $\forall x\, \Box(A(x))$ as an abbreviation for the arithmetical sentence which formalizes the fact that for all x there is a $I\Delta_0 + \Omega_1$-proof of $A(\dot{x})$, where $\dot{x}$ is the binary numeral for x. If A and B are arithmetical sentences, $\Box A \leq \Box B$ denotes the witness comparison sentence

$$\exists x\, (\mathrm{Prf}_{I\Delta_0+\Omega_1}(x, \ulcorner A \urcorner) \wedge \forall y < x\, \neg \mathrm{Prf}_{I\Delta_0+\Omega_1}(y, \ulcorner B \urcorner)).$$

Similarly $\Box A < \Box B$ denotes

$$\exists x\, (\mathrm{Prf}_{I\Delta_0+\Omega_1}(x, \ulcorner A \urcorner) \wedge \forall y \leq x\, \neg \mathrm{Prf}_{I\Delta_0+\Omega_1}(y, \ulcorner B \urcorner)).$$

$\Box_k A$ is a formalization of the fact that A has a proof in $I\Delta_0 + \Omega_1$ of Gödel number $\leq k$. So $\Box A < \Box B$ can be written as $\exists x\, (\Box_x A \wedge \neg\Box_x B)$. (Note that all the above definitions are only abbreviations for some arithmetical formulas and are not meant to correspond to an enrichment of the modal language.)

Remark 2.3. Since the proof predicate can be formalized by a Σ_1^b-formula, we have $I\Delta_0 + \Omega_1 \vdash \Box A \rightarrow \Box\Box A$ and $I\Delta_0 + \Omega_1 \vdash \Box_x A \rightarrow \Box\Box_x A$.

Definition 2.4. By an $I\Delta_0 + \Omega_1$-*cut* we mean a formula $I(x)$ with exactly one free variable x, such that $I\Delta_0 + \Omega_1$ proves that I defines an initial segment of numbers containing 0 and closed under successor, addition, multiplication, and the function ω_1 (see [15]). We write $x \in I$ for $I(x)$.

Given an $I\Delta_0 + \Omega_1$-cut I, $I\Delta_0 + \Omega_1$ can formalize the fact that I defines a model of $I\Delta_0 + \Omega_1$. It follows that for any arithmetical sentence θ we have:

Proposition 2.5. $I\Delta_0 + \Omega_1 \vdash \Box(\theta) \rightarrow \Box(\theta^I)$, *where* θ^I *is obtained from* θ *by relativizing all the quantifiers to* I.

Note that if a Σ_1^0-formula is witnessed in a cut, then it is witnessed in the universe. Thus we have:

Remark 2.6. For every $I\Delta_0 + \Omega_1$-cut I, and every Σ_1^0-formula $\sigma(x_1, \ldots, x_n)$,

$$I\Delta_0 + \Omega_1 \vdash x_1 \in I \wedge \cdots \wedge x_n \in I \wedge \sigma^I(x_1, \ldots, x_n) \rightarrow \sigma(x_1, \ldots, x_n).$$

The use of binary numerals is essential for the following proposition (see [7]):

Proposition 2.7. *For any* $I\Delta_0 + \Omega_1$*-cut* I, $I\Delta_0 + \Omega_1 \vdash \forall x\, \Box(x \in I)$.

Making use of an efficient truth predicate (as in [7]), Verbrugge [13] proved the following result:

Theorem 2.8 (Small reflection principle). $I\Delta_0 + \Omega_1 \vdash \forall k\, \Box(\Box_k A \rightarrow A)$.

An immediate corollary is the following principle (originally stated by Švejdar for PA):

Corollary 2.9 (Švejdar's principle). $I\Delta_0 + \Omega_1 \vdash \Box A \rightarrow \Box(\Box B \leqslant \Box A \rightarrow B)$.

Using Solovay's technique of shortening of cuts, it is easy to prove the following:

Proposition 2.10. *There is an* $I\Delta_0 + \Omega_1$*-cut* J, *such that for each* Σ_1^0*-formula* $\sigma(x_1, \ldots, x_n)$ *we have*:

$$I\Delta_0 + \Omega_1 \vdash J(x_1) \wedge \cdots \wedge J(x_n) \wedge \sigma^J(x_1, \ldots, x_n) \rightarrow \Box\sigma(x_1, \ldots, x_n).$$

Proof. The proof is similar to the proof of provable Σ_1^b-completeness for $I\Delta_0 + \Omega_1$ (see [15]). Therefore we only give a sketch of the proof. By induction on the structure of the formula, one can prove that for each Δ_0-formula A with free variables $x_1, \ldots, x_n$, there are k, l and m such that

$$I\Delta_0 + \Omega_1 \vdash \forall x_1, \ldots, x_n\, \forall x\, \forall y\, (x = \max(x_1, \ldots, x_n) \wedge |y| = 2^{|\ulcorner A \urcorner|^k \cdot |x|^l} + m$$
$$\wedge A(x_1, \ldots, x_n) \rightarrow \exists z \leqslant y\, \mathrm{Prf}_{I\Delta_0 + \Omega_1}(z, \ulcorner A(\dot{x}_1, \ldots, \dot{x}_n) \urcorner)).$$

Now let J be the cut, which can be obtained by Solovay's shortening methods (cf. [15, 8, 10]), such that

- $I\Delta_0 + \Omega_1 \vdash \forall x\, (J(x) \rightarrow \exists z\, (z = 2^x))$ and
- $I\Delta_0 + \Omega_1 \vdash \forall x, y\, (J(x) \wedge J(y) \rightarrow J(x + y) \wedge J(x \cdot y) \wedge J(2^{|x| \cdot |y|}))$.

For this cut, we have for all Δ_0-formulas A,

$$I\Delta_0 + \Omega_1 \vdash \forall x_1, \ldots, x_n\, (J(x_1) \wedge \cdots \wedge J(x_n) \wedge A(x_1, \ldots, x_n) \\ \rightarrow \exists z\, \mathrm{Prf}_{I\Delta_0+\Omega_1}(z, \ulcorner A(\dot{x}_1, \ldots, \dot{x}_n)\urcorner)).$$

The result immediately follows. □

In the sequel 'J' will always refer to the cut of Proposition 2.10.

Corollary 2.11. *If S_i $(i = 1, \ldots, k)$ are Σ_1^0-sentences, then*

$$I\Delta_0 + \Omega_1 \vdash \Box\Big(\bigvee_i S_i\Big) \rightarrow \Box\Big(\bigvee_i \Box^+ S_i\Big).$$

Proof. Let J be as in Proposition 2.10. Work in $I\Delta_0 + \Omega_1$ and suppose $\Box(\bigvee_i S_i)$ holds. Since J (provably) defines a model of $I\Delta_0 + \Omega_1$, it follows $\Box(\bigvee_i S_i^J)$. By Proposition 2.10 and Remark 2.6, $\Box(S_i^J \rightarrow \Box^+ S_i)$ and the desired result follows. □

The above corollary was originally proved by Visser [14] as a consequence of the following more general result:

Theorem 2.12 (Visser's principle). *If S and S_i $(i = 1, \ldots, k)$ are Σ_1^0-sentences, then*

$$I\Delta_0 + \Omega_1 \vdash \Box\Big(\bigwedge_i (S_i \rightarrow \Box S_i) \rightarrow S\Big) \rightarrow \Box S.$$

3. Trees of undecidable sentences

We will rephrase the problem of whether $PL\Omega = L$ as a problem concerning the existence of suitable trees of undecidable sentences.

Let $\mathscr{C}$ be a class of finite tree-like strict partial orders. Without loss of generality we assume that for all $(K, <) \in \mathscr{C}$, $K = \{1, \ldots, n\}$ for some $n \in \omega$, and 1 is the root (i.e., the least element of K). By $Th(\mathscr{C})$ we denote the set of all those modal formulas that are forced at the root of every Kripke model whose underlying tree belongs to $\mathscr{C}$. Let $\leqslant$ be the non-strict partial order associated to $<$.

Definition 3.1. Given a tree $(K, <)$ with root 1 and underlying set $K = \{1, \ldots, n\}$, we say that $(K, <)$ can be *embedded* (or *simulated*) in $I\Delta_0 + \Omega_1$ if there are arithmetical sentences $L_1, \ldots, L_n$ (one for each node) such that, letting □ denote formalized provability from $I\Delta_0 + \Omega_1$, the conjunction of the following sentences is consistent with $I\Delta_0 + \Omega_1$:

1. L_1;
2. $\Box^+(L_1 \vee \cdots \vee L_n)$;
3. $\Box^+(L_i \to \neg L_j)$ for $i \neq j$ in K;
4. $\Box^+(L_a \to \Diamond L_b)$ for $a < b$ in K;
5. $\Box^+(L_a \to \Box\neg L_b)$ for $a \not< b$ in K.

The following lemma is inspired by Solovay's proof of the fact that L is the provability logic of PA.

Lemma 3.2. *In order for $PL\Omega \subseteq Th(\mathscr{C})$ to be the case it suffices that every tree $(K, <) \in \mathscr{C}$ can be embedded in $I\Delta_0 + \Omega_1$.*

Proof. Suppose $A \notin Th(\mathscr{C})$. Then there is a Kripke model $(K, <, \Vdash)$ such that $(K, <) \in \mathscr{C}$, $K = \{1, \ldots, n\}$, 1 is the least element of K, and $1 \Vdash \neg A$. By our hypothesis there exists a model M of $I\Delta_0 + \Omega_1$ and sentences $L_1, \ldots, L_n$ satisfying, inside the model M, the properties 1–5 of Definition 3.1. Define an $I\Delta_0 + \Omega_1$-interpretation $*$ by setting, for every atomic propositional letter p, $p^* \equiv \bigvee_{i \Vdash p} L_i$. It is then easy to verify by induction on the complexity of the modal formula B, that for every $i \in K$:

1. $i \Vdash B \Rightarrow M \models \Box^+(L_i \to B^*)$;
2. $i \Vdash \neg B \Rightarrow M \models \Box^+(L_i \to \neg B^*)$.

The induction step for $\Box$ is based on the following consequences of 1–5:

1. $M \models \Box^+(L_i \to \Diamond L_j)$ for $i < j$;
2. $M \models \Box^+\left(L_i \to \Box\left(\bigvee_{j>i} L_j\right)\right)$.

Since $1 \Vdash \neg A$, it follows that $M \models \neg A^*$, hence $I\Delta_0 + \Omega_1 \nvdash A^*$ as desired. $\Box$

Corollary 3.3. *If every finite tree $(K, <)$ can be embedded in $I\Delta_0 + \Omega_1$, then $PL\Omega = L$.*

Proof. Let $\mathscr{C}$ be the class of all finite trees. If our hypothesis is satisfied, then $L \subseteq PL\Omega \subseteq Th(\mathscr{C}) = L$. $\Box$

It can be easily verified that the sufficient condition of Lemma 3.2 is also necessary. Thus $PL\Omega \subseteq Th(\mathscr{C})$ iff every $(K, <) \in \mathscr{C}$ can be embedded in $I\Delta_0 + \Omega_1$. Hence a very natural question to ask is:

Question 3.4. Which finite trees can be embedded in $I\Delta_0 + \Omega_1$?

Note that a complete answer to the above question, although interesting by itself, may not suffice to characterize $PL\Omega$. In fact if $\mathscr{C}$ is the set of *all* finite trees

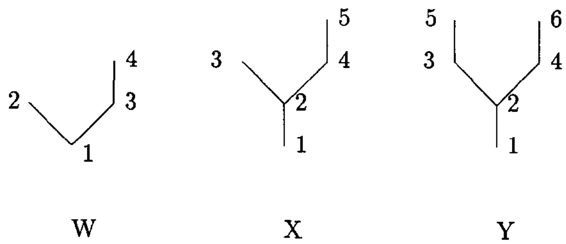

Fig. 1. The trees **W**, **X**, **Y**.

that can be embedded in $I\Delta_0 + \Omega_1$, we can in general only conclude $PL\Omega \subseteq Th(\mathscr{C})$.

In order to describe the results proved in this and previous papers, we need to define what it means for a tree to omit another tree.

Definition 3.5. Let $(T_1, <_1)$ and $(T_2, <_2)$ be (strict) partial orders. An homomorphic embedding of $(T_1, <_1)$ into $(T_2, <_2)$ is an injective map $f: T_1 \to T_2$ such that for all $x, y \in T_1$, $x <_1 y \leftrightarrow f(x) <_2 f(y)$. If there is no homomorphic embedding of T_1 into T_2 we say that T_2 *omits* T_1.

If we try to adapt Solovay's proof to $I\Delta_0 + \Omega_1$ in the most straightforward manner, the only trees that we can embed in $I\Delta_0 + \Omega_1$ are the linear trees, namely trees omitting $(K, <)$ where $K = \{1, 2, 3\}$, $1 < 2$, $1 < 3$ and 2 is incomparable with 3.

A first improvement can be achieved using Švejdar's principle: let $\mathscr{C}_1$ be the class of all trees that omit the tree $\mathbf{W} = (W, <)$, the least strict partial order with underlying set $W = \{1, 2, 3, 4\}$ such that $1 < 2$, $1 < 3 < 4$ (see Fig. 1). The second author proved in her master's thesis [12] that for trees in $\mathscr{C}_1$ Solovay's proof can be adapted using Švejdar's principle. In other words, $PL\Omega \subseteq Th(\mathscr{C}_1)$. She also proved that the inclusion is a strict one.

In subsequent work she showed, using both Švejdar's and Visser's principles, that $PL\Omega$ is included in the modal theory of $\mathscr{C}_2$, the class of all trees of height $\leqslant 3$.

A new improvement [2] was achieved by analogous techniques but using a different definition of the Solovay constants. In this way it was proved that $PL\Omega \subseteq Th(\mathscr{C}_3)$, where $\mathscr{C}_3$ is the class of all trees that omit the tree $\mathbf{X} = (X, <)$, the least strict partial order with underlying set $X = \{1, 2, 3, 4, 5\}$ such that $1 < 2 < 4 < 5$, $1 < 2 < 3$.

Finally in Section 4 of the present paper, we improve these earlier results, by proving:

Theorem 3.6. *$PL\Omega \subseteq Th(\mathscr{C}_4)$, where $\mathscr{C}_4$ is the class of trees that omit the tree $\mathbf{Y} = (Y, <)$, the least strict partial order with underlying set $Y = \{1, 2, 3, 4, 5, 6\}$ such that* $1 < 2 < 3 < 5$, $1 < 2 < 4 < 6$.

In particular, Theorem 3.6 says that we can embed **X** but not **Y**. Note that the trees in $\mathscr{C}_4$ can have an arbitrarily large number of bifurcation points, but each bifurcation point except the root can have at most one immediate successor which is not a leaf. The root can have any number of immediate successors which are not leaves.

On the other hand, we prove in Sections 5 and 6 that for many classes $\mathscr{C}$ of trees (and especially for the classes $\mathscr{C}_1, \ldots, \mathscr{C}_4$ defined above), we cannot have $PL\Omega = Th(\mathscr{C})$. Therefore, all inclusions mentioned above are strict. More precisely we prove that *if* $PL\Omega = Th(\mathscr{C})$, then every binary tree can be homomorphically embedded in some tree belonging to $\mathscr{C}$. So it is unlikely that $PL\Omega$ is the theory of a class of trees, unless $PL\Omega = L$.

4. Upper bounds on $PL\Omega$

Our task in this section will be to prove $PL\Omega \subseteq Th(\mathscr{C}_4)$ using Lemma 3.2.

Definition 4.1. Given $(K, <) \in \mathscr{C}_4$, we say that $i \in K$ is a *special* node, iff i is a leaf, and some brother of i is not a leaf.

For example, in the tree **X** of Fig. 1, the only special node is 3.

Definition 4.2. Let $(K, <) \in \mathscr{C}_4$. Without loss of generality assume that $K = \{1, \ldots, n\}$ and 1 is the root. Let J be the cut of Proposition 2.10. By a self-referential construction based on the diagonal lemma, we can simultaneously define sentences $L_1, \ldots, L_n$, and auxiliary functions v, w, S, such that the following holds:

1. If $i \in K$ is not special, let $w(i) = \mu x\, \Box_x \neg L_i$ (with the convention that $w(i) = \infty$ if $\Diamond L_i$); if $i \in K$ is special $w(i) = \mu x \in J\, \Box_x \neg L_i$ (with the convention that $w(i) = \infty$ if $\Diamond^J L_i$). We agree that ∞ is a specific element greater than any integer. Note that the definition of w can be formalized in $I\Delta_0 + \Omega_1$.
2. If j is an immediate successor of i in $(K, <)$, let $v(i, j) = w(j)$; otherwise $v(i, j) = \infty$.
3. $S: K \to K$ is defined as follows: $S(i) = i$ if for no $j \in K$ we have $v(i, j) < \infty$; otherwise among all the $j \in K$ with $v(i, j) < \infty$, pick one for which $v(i, j)$ is minimal, and set $S(i) = S(j)$. (Note that there exists at most one such j because if $w(j) = w(j') < \infty$, then there is one single proof of both $\neg L_j$ and $\neg L_{j'}$, so $j = j'$.)
4. $I\Delta_0 + \Omega_1 \vdash L_i \leftrightarrow \Box\neg L_1 \wedge i = S(1)$.

The important point to observe, is that the definition of S can be formalized in $I\Delta_0 + \Omega_1$ and that $I\Delta_0 + \Omega_1$ proves that $S(1)$ is always defined. This depends on the fact that, although S is defined in a recursive way, to compute $S(1)$ one only needs a standard number of recursive calls, namely at most d where d is the height of the tree $(K, \prec)$ (in fact at each recursive call we climb one step up in the tree). Note also that S depends self-referentially on $L_1, \ldots, L_n$. Finally note that, if a, b are distinct immediate successors of i, then the statement $v(i, a) < v(i, b)$ is equivalent to a witness comparison sentence in which some quantifiers are relativized to J. In particular, if a and b are not special, then $v(i, a) < v(i, b)$ is equivalent to the Σ_1^0-sentence $\Box\neg L_a < \Box\neg L_b$.

Remark 4.3. The main differences with Solovay's construction are the following: (1) We do not use an extra node 0 (but this is a minor point since we could define L_0 as $\Diamond L_1$). (2) In our construction we can only jump one step at a time, namely at each recursive call S we can only move from one point to some immediate successor. (3) While Solovay employs a primitive recursive function from ω to K whose definition is not directly formalizable in $I\Delta_0 + \Omega_1$, we use instead a function $S: K \to K$ which is provably total in $I\Delta_0 + \Omega_1$. (4) We jump to a special node $i \in K$ only if we find a proof of $\neg L_i$ belonging to the cut J.

Given $(K, \prec)$ as above, we will show that $L_1, \ldots, L_n$ constitute an embedding of $(K, \prec)$ in $I\Delta_0 + \Omega_1$. We need the following lemma.

Lemma 4.4. *Let $L_1, \ldots, L_n$ and $(K, \prec)$ be as in Definition* 4.2. *Then*:

1. $\vdash \Box\neg L_1 \to L_1 \vee \cdots \vee L_n$.
2. $\vdash L_i \to \neg L_j$ *for* $i \neq j$ *in* K.
3. $\vdash L_i \to \Box\neg L_i$ *for* $i \in K$.
4. L_1 *is consistent with* $I\Delta_0 + \Omega_1$.
5. *If* $j, j' \in K$ *are brothers, then* $\vdash \Box\neg L_j \leftrightarrow \Box\neg L_{j'}$.
6. $\vdash L_a \to \Diamond L_b$ *for* $a \prec b$ *in* K.
7. $\vdash L_b \to \Box\neg L_a$ *for* $a \prec b$ *in* K.
8. *If i is above* (*i.e.* $\succcurlyeq$) *a brother of j, then* $\vdash L_i \to \Box\neg L_j$; *if moreover j is a leaf, then* $\vdash L_j \to \Box\neg L_i$.
9. *Let $b \succ 1$ be an immediate successor of the root* 1. *Then* $\vdash L_1 \to \Box\Box(\neg L_b)$.
10. $\vdash L_1 \to \Box^+(L_i \to \Box\neg L_j)$ *whenever i, j are incomparable nodes of K.*

Here '$\vdash$' *stands for* '$I\Delta_0 + \Omega_1 \vdash$'.

Proof. It will be clear from the context at which places we reason inside $I\Delta_0 + \Omega_1$.

(1) and (2) are clear from the definition of the sentences L_i and the fact that $S: K \to K$ is a total function.

(3) L_i implies that $\Box\neg L_1 \wedge i = S(1)$. If $i = 1$, $\Box\neg L_i$ follows immediately; otherwise we have $w(i) < \infty$, and therefore $\Box\neg L_i$.

(4) If L_1 is inconsistent with $I\Delta_0 + \Omega_1$, then $\Box\neg L_1$ holds in the standard model, so by (1), one of the sentences L_i must hold in the standard model. This is absurd since each of these sentences implies its own inconsistency.

(5) First note that $\vdash \Box_x\neg L_j \to \Box(x \in J \wedge \Box_x\neg L_j)$. Thus, regardless of whether j is special or not, $\vdash \Box\neg L_j \to \Box(w(j) = \mu x\, \Box_x\neg L_j)$. Since j and j' are brothers,

$\vdash L_{j'} \to w(j') < w(j)$ (because $j' = S(1)$ implies $w(j') < w(j)$). Therefore $\vdash \Box\neg L_j \to \Box(L_{j'} \to \Box\neg L_{j'} < \Box\neg L_j)$. On the other hand by Švejdar's principle $\vdash \Box\neg L_j \to \Box(\Box\neg L_{j'} < \Box\neg L_j \to \neg L_{j'})$ and we can conclude $\vdash \Box\neg L_j \to \Box\neg L_{j'}$.

(6) In $I\Delta_0 + \Omega_1$ we can formalize the fact that if a consistent theory proves the consistency of another theory, then the latter is consistent (we assume that all theories contain $I\Delta_0 + \Omega_1$ and have a Σ_1^b set of axioms). Hence $\vdash \Diamond L_u \wedge \Box(L_u \to \Diamond L_v) \to \Diamond L_v$. It follows that in the proof of (6) we can assume without loss of generality that b is an immediate successor of a. Working inside $I\Delta_0 + \Omega_1$, assume L_a. Then $a = S(1)$. Hence $w(b) = \infty$. Now if b is not a special node, then $w(b) = \infty \leftrightarrow \Diamond L_b$ and we are done. If b is a special node, from $w(b) = \infty$ we can only conclude $\Diamond^J L_b$, so we need an additional argument. This is provided by point (5). In fact by definition of special node, a has certainly one immediate successor b' which is not special. Hence from L_a we can derive $\Diamond L_{b'}$ reasoning as above. By point (5), $\Diamond L_b \leftrightarrow \Diamond L_{b'}$ and we are done.

(7) can be derived through the chain of implications: $L_b \to \Box\neg L_b \to \Box\Box\neg L_b \to \Box\neg L_a$, where the last implication uses point (6).

(8) Let i be above a brother of j. Then by (5), (7) and (3) $\vdash L_i \to \Box\neg L_j$ as desired. To prove the second part, assume further that j is a leaf. We need to show $\vdash L_j \to \Box\neg L_i$. We can assume that i is *strictly* above a brother j' of j (for if i itself is a brother of j the desired result follows from (3) an (5)). But then j must be a special node, and therefore $w(j) = \mu x \in J\, \Box_x \neg L_j$. So $w(j) < w(j')$ is equivalent to a Σ_1^0-formula relativized to J, namely

$$w(j) < w(j') \leftrightarrow \exists x \in J\, (\mathrm{Prf}_{I\Delta_0+\Omega_1}(x, \ulcorner\neg L_j\urcorner) \wedge \forall y \leqslant x\, \neg\mathrm{Prf}_{I\Delta_0+\Omega_1}(y, \ulcorner\neg L_{j'}\urcorner)).$$

Thus by the properties of the cut J (and by Theorem 2.7), $\vdash w(j) < w(j') \to \Box w(j) < w(j')$. Now the desired result follows by observing that $\vdash L_j \to w(j) < w(j')$ (as $\vdash j = S(1) \to w(j) < w(j')$) and $\vdash L_i \to w(j') < w(j)$.

(9) By (1) and (3), $\vdash L_1 \to \Box(\bigvee_{i>1} L_i)$. So to prove $\vdash L_1 \to \Box\Box\neg L_b$, it suffices to show that for each $i > 1$ we have $\vdash \Box(L_i \to \Box\neg L_b)$. This follows from (8), (3) and (7).

(10) If the incomparable nodes i and j are in one of the situations covered by point (8), then $\vdash L_i \to \Box\neg L_j$, and a fortiori $\vdash L_1 \to \Box^+(L_i \to \Box\neg L_j)$ as desired. Since $(K, <)$ omits $\mathbf{Y}$, (8) can always be applied except when the biggest node (with respect to $\leqslant$) below i and j is 1 (the root). So assume that this is the case. By (2), we have $\vdash L_1 \to (L_i \to \Box\neg L_j)$. In order to show that also $\vdash L_1 \to \Box(L_i \to \Box\neg L_j)$, we will make use of Proposition 2.10. Let i', j' be the least nodes with $1 < i' \leqslant i$ and $1 < j' \leqslant j$. So i' and j' are brothers. It follows from (9) that $\vdash L_1 \to \Box(\Box\neg L_{i'})$. Therefore, by Proposition 2.5, $\vdash L_1 \to \Box(\Box^J \neg L_{i'})$. In the presence of $\Box^J\neg L_{i'}$, the sentence $w(i') < w(j')$ is equivalent to a Σ_1^0-sentence relativized to J. Therefore, by Proposition 2.10, $\vdash L_1 \to \Box(w(i') < w(j') \to \Box(w(i') < w(j')))$. The desired result now follows from the fact that L_i provably implies $i = S(1)$ which entails $w(i') < w(j')$, while L_j provably implies $w(j') < w(i')$. $\Box$

Corollary 4.5. *If* $(K, <)$ *and* $L_1, \ldots, L_n$ *are as above, then the conjunction of the following sentences is consistent with* $I\Delta_0 + \Omega_1$:

1. L_1;
2. $\Box^+(L_1 \vee \cdots \vee L_n)$;
3. $\Box^+(L_i \to \neg L_j)$ *for* $i \neq j$ *in* K;
4. $\Box^+(L_a \to \Diamond L_b)$ *for* $a < b$ *in* K;
5. $\Box^+(L_a \to \Box\neg L_b)$ *for* a $a \nless b$ *in* K.

Proof. The derivation of Corollary 4.5 from Lemma 4.4 follows from a straightforward argument which can even be formalized in the decidable theory L^ω. (The axioms of L^ω are all the theorems of L and all the instances of $\Box A \to A$. The only rule is modus ponens.) $\Box$

We have thus shown that every tree of $\mathscr{C}_4$ can be embedded in $I\Delta_0 + \Omega_1$. Thus:

Theorem 4.6. $PL\Omega \subseteq Th(\mathscr{C}_4)$.

5. Disjunction property

In this section we prove the following:

Theorem 5.1. *IF* $PL\Omega = Th(\mathscr{C})$, *where* $\mathscr{C}$ *is a class of finite trees, then every binary tree can be homomorphically embedded in some tree belonging to* $\mathscr{C}$.

In particular, since the binary tree **Y** cannot be embedded in any member of $\mathscr{C}_4$, it will follow that the inclusion $PL\Omega \subseteq Th(\mathscr{C}_4)$ is strict.

We will use the fact that $PL\Omega$ has the 'disjunction property' as proved by Franco Montagna (private communication).

Definition 5.2. A modal theory P has the disjunction property if for every pair of modal sentences A and B, if $P \vdash \Box A \vee \Box B$, then $P \vdash A$ or $P \vdash B$.

It is known that L has the disjunction property.

Theorem 5.3 (Montagna). *$PL\Omega$ has the disjunction property.*

Proof. Suppose that for some $I\Delta_0 + \Omega_1$-interpretations $^\circ$ and $^\bullet$ we have $I\Delta_0 + \Omega_1 \nvdash A(\boldsymbol{p}^\circ)$ and $I\Delta_0 + \Omega_1 \nvdash B(\boldsymbol{p}^\bullet)$, where $\boldsymbol{p}$ contains all propositional variables occurring in the modal formulas A and B. We have to prove that there is an $I\Delta_0 + \Omega_1$-interpretation * such that $I\Delta_0 + \Omega_1 \nvdash (\Box A \vee \Box B)^*$.

By multiple diagonalization, define for all $p_i \in \boldsymbol{p}$ an arithmetical formula p_i^* such that

$$I\Delta_0 + \Omega_1 \vdash p_i^* \leftrightarrow (\Box A(\boldsymbol{p}^*) \leqslant \Box B(\boldsymbol{p}^*) \wedge p_i^\circ) \vee (\Box B(\boldsymbol{p}^*) < \Box A(\boldsymbol{p}^*) \wedge p_i^\bullet).$$

We will show that $I\Delta_0 + \Omega_1 \nvdash (\Box A \vee \Box B)^*$. So suppose, to derive a contradiction, that $I\Delta_0 + \Omega_1 \vdash \Box A(\boldsymbol{p}^*) \vee \Box B(\boldsymbol{p}^*)$. Then

$$I\Delta_0 + \Omega_1 \vdash \Box A(\boldsymbol{p}^*) \leqslant \Box B(\boldsymbol{p}^*) \vee \Box B(\boldsymbol{p}^*) < \Box A(\boldsymbol{p}^*).$$

Thus, because $I\Delta_0 + \Omega_1$ is a true theory, either

1. $\Box A(\boldsymbol{p}^*) \leqslant \Box B(\boldsymbol{p}^*)$ and $I\Delta_0 + \Omega_1 \vdash p_i^* \leftrightarrow p_i^\circ$ for all i

(by definition of $\boldsymbol{p}^*$), or

2. $\Box B(\boldsymbol{p}^*) < \Box A(\boldsymbol{p}^*)$ and $I\Delta_0 + \Omega_1 \vdash p_i^* \leftrightarrow p_i^\bullet$ for all i.

In case 1, we have $I\Delta_0 + \Omega_1 \vdash A(\boldsymbol{p}^*)$, so $I\Delta_0 + \Omega_1 \vdash A(\boldsymbol{p}^\circ)$, contradicting our assumption. Similarly, case 2 contradicts the assumption $I\Delta_0 + \Omega_1 \nvdash B(\boldsymbol{p}^\bullet)$. □

In order to prove Theorem 5.1 we need the following definition.

Definition 5.4. We define D_n by induction.
- $D_0 = \top$.
- $D_{i+1}(\boldsymbol{p}, r) = \Diamond(D_i(\boldsymbol{p}) \wedge \Box^+ r) \wedge \Diamond(D_i(\boldsymbol{p}) \wedge \Box^+ \neg r)$, where $\boldsymbol{p}$ is of length i, and all propositional variables in $\boldsymbol{p}$, r are different.

The main property of the formulas D_n is expressed by the following lemma.

Lemma 5.5. *If* $\mathbf{K}$ *is a finite tree-like Kripke model with root* k *such that* $k \Vdash D_n$, *then we can homomorphically embed* (*see Definition* 3.5) *the full binary tree* $\mathbf{T}_n$ *of* $2^{n+1} - 1$ *nodes into* $\mathbf{K}$.

Proof. By induction on n.

Base case. Trivial: $\mathbf{T}_0$ contains only one point.

Induction step. Suppose that $k \Vdash D_{i+1}(\boldsymbol{p}, r)$, i.e.,

$$k \Vdash \Diamond(D_i(\boldsymbol{p}) \wedge \Box^+ r) \wedge \Box(D_i(\boldsymbol{p}) \wedge \Box^+ \neg r).$$

Then there are nodes k_1, k_2 such that $k \preccurlyeq k_1$, $k \preccurlyeq k_2$, $k_1 \Vdash D_i(\boldsymbol{p}) \wedge \Box^+ r$ and $k_2 \Vdash D_i(\boldsymbol{p}) \wedge \Box^+ \neg r$. By the induction hypothesis, we can homomorphically embed a copy of the full binary tree $\mathbf{T}_i$ of bifurcation depth i into the subtree of $\mathbf{K}$ that consists of all points $\succcurlyeq k_1$. Analogously, we can homomorphically embed a copy of $\mathbf{T}_i$ into the subtree of $\mathbf{K}$ of points $\succcurlyeq k_2$.

Because $k_1 \Vdash \Box^+ r$ and $k_2 \Vdash \Box^+ \neg r$, we may conclude that k_1 and k_2 are incomparable and that the two images of $\mathbf{T}_i$ are disjoint. Therefore, we can combine both homomorphic embeddings into one and subsequently map the root of $\mathbf{T}_{i+1}$ to k. Thus an homomorphic embedding of $\mathbf{T}_{i+1}$ into $\mathbf{K}$ is produced. □

Theorem 5.1 is now an immediate consequence of the following:

Theorem 5.6. *Let $\mathscr{C}$ be a class of finite trees such that $Th(\mathscr{C})$ has the disjunction property. Then for every n, $Th(\mathscr{C}) + D_n$ is consistent. Thus every binary tree can be homomorphically embedded in some member of $\mathscr{C}$.*

Proof. Let $P = Th(\mathscr{C})$. Note that $P \supseteq L$. We prove by induction on n that $P + D_n$ is consistent.

Base case. Trivial.

Induction step. Suppose as induction hypothesis that for any $\boldsymbol{p}$ consisting of i different propositional variables, $P + D_i(\boldsymbol{p})$ is consistent. In order to derive a contradiction, suppose that $P \vdash \neg D_{i+1}(\boldsymbol{p}, r)$, that is

$$P \vdash \Box(\neg D_i(\boldsymbol{p}) \vee \neg\Box^+ r) \vee \Box(\neg D_i(\boldsymbol{p}) \vee \neg\Box^+\neg r).$$

Then by the disjunction property, either

1. $P \vdash \neg D_i(\boldsymbol{p}) \vee \neg\Box^+ r$ or
2. $P \vdash \neg D_i(\boldsymbol{p}) \vee \neg\Box^+\neg r$.

We show that 1 cannot hold. By the induction hypothesis, $P \nvdash \neg D_i(\boldsymbol{p})$. Since r does not appear in $D_i(\boldsymbol{p})$, we can take $r = \top$. But then $P \vdash \Box^+ r$, so $P \nvdash \neg D_i(\boldsymbol{p}) \vee \neg\Box^+ r$.

By an analogous proof, we can show that 2 cannot hold, which gives the desired contradiction. □

Note that in the proof of the fact that $Th(\mathscr{C}) + D_n$ is consistent we have only used the fact that $Th(\mathscr{C})$ is a consistent modal theory extending L and satisfying the disjunction property. The same proof can therefore be applied to $PL\Omega$, yielding:

Proposition 5.7. *$PL\Omega + D_n$ is consistent.*

We are now able to strengthen Theorem 5.1 as follows:

Theorem 5.8. *If there exists a binary tree H which cannot be homomorphically embedded in any member of $\mathscr{C}$, then $Th(\mathscr{C}) \nsubseteq PL\Omega$.*

Proof. Under our assumption there is some n such that the full binary tree of height n cannot be embedded in any member of $\mathscr{C}$. Hence $Th(\mathscr{C}) + D_n$ is inconsistent. On the other hand $PL\Omega + D_n$ is consistent. □

6. Further results

We give some further results, due to the first author, of the form '$PL\Omega + \phi$ is consistent', for various choices of ϕ. In particular we strengthen Proposition 5.7

by showing that $PL\Omega + D_n + \Box^{n+1}\bot$ is consistent. Note, for a motivation, that $L = PL\Omega$ if and only if every modal formula ϕ consistent with L, is consistent with $PL\Omega$. The disjunction property will not be used.

Definition 6.1. Given a tree $(K, <)$ with root 1 and underlying set $K = \{1, \ldots, n\}$, we say that $(K, <)$ can be *weakly embedded* in $I\Delta_0 + \Omega_1$ if there are arithmetical sentences $L_1, \ldots, L_n$ (one for each node) such that, letting $\Box$ denote formalized provability from $I\Delta_0 + \Omega_1$, the conjunction of the following sentences is consistent with $I\Delta_0 + \Omega_1$:

1. L_1;
2. $\Box^+(L_i \to \neg L_j)$ for $i \neq j$ in K;
3. $\Box^m\bot \wedge \neg\Box^{m-1}\bot$ where m is the height of $(K, <)$ (i.e., the maximum cardinality of a chain in $(K, <)$). We agree that $\Box^0\bot$ is $\bot$ and $\Box^{k+1}\bot$ is $\Box\Box^k\bot$;
4. $\Box^+(L_a \to \Diamond L_b)$ for $a < b$ in K;
5. $\Box^+(L_a \to \Box\neg L_b)$ for $a \nless b$ in K.

It is easy to verify that 'embeddable' implies 'weakly embeddable'. (The only point to check is 3.) We will prove:

Theorem 6.2. *Every finite tree K can be weakly embedded in $I\Delta_0 + \Omega_1$.*

This is to be compared with the previous result Theorem 3.6 saying that every tree omitting **Y** can be (strongly) embedded in $I\Delta_0 + \Omega_1$.

Note that the fact that K is weakly embeddable in $I\Delta_0 + \Omega_1$ can be expressed in the form '$PL\Omega + \phi_K$ is consistent', where ϕ_K is a suitable modal formula depending on K (i.e., the conjunction of the five sentences of Definition 6.1, where the L_i's are now thought as atomic modal formulas).

Corollary 6.3. *$PL\Omega + D_n + \Box^{n+1}\bot$ is consistent.*

The proof of the corollary is easy and left to the reader. The idea is that the arithmetical sentences needed to prove that $PL\Omega + D_n + \Box^{n+1}\bot$ is consistent, can be obtained as boolean combinations of the sentences L_i which weakly embed the full binary tree of height $n + 1$ in $I\Delta_0 + \Omega_1$.

Theorem 6.2 will be proved with the help of a self-referential construction based on an auxiliary tree $K_1 \supseteq K$ which is obtained by duplicating each bifurcation node of K. The idea is that we can do in two steps what we cannot do in one step.

Definition 6.4. Given a finite tree $(K, <)$, we injectively associate, to each bifurcation node i of $(K, <)$, a new node $d(i)$ not in K, and we define K_1 as K union the set of all the new nodes $d(i)$. We make K_1 into a tree $(K_1, <_1)$ by putting each $d(i)$ immediately above i and by stipulating that the immediate successors of $d(i)$ in $(K_1, <_1)$ are the immediate successors of i in $(K, <)$. Briefly: $(K_1, <_1)$ is obtained from $(K, <)$ by duplicating each bifurcation node.

On a first reading of the rest of this section we suggest to think of $(K, <)$ as the tree **Y** of Fig. 1.

Definition 6.5. Let J be the cut of Proposition 2.10. Let $(K_1, <_1)$ be obtained from $(K, <)$ by duplicating each bifurcation node. By the diagonal lemma, we simultaneously define sentences L_i for $i \in K_1$, and auxiliary functions v, w, S such that the following holds:

1. If $j \in K_1$ is an immediate successor of one of the new nodes $d(i) \in K_1 - K$, then $w(j) = \mu x \in J\ (\Box_x \neg L_j \wedge \Diamond_x L_{d(i)})$; otherwise $w(j) = \mu x\ \Box_x \neg L_j$.
2. If $j \in K_1$ is an immediate successor of i in $(K_1, <_1)$, let $v(i, j) = w(j)$; otherwise $v(i, j) = \infty$.
3. $S : K_1 \to K_1$ is defined as follows: $S(i) = i$ if for no $j \in K_1$ we have $v(i, j) < \infty$; otherwise among all the $j \in K_1$ with $v(i, j) < \infty$, pick one for which $v(i, j)$ is minimal, and set $S(i) = S(j)$. (Note that there exists at most one such j.)
4. For $i \in K_1$, $I\Delta_0 + \Omega_1 \vdash L_i \leftrightarrow \Box\neg L_1 \wedge i = S(1)$.

Remark 6.6. Note that the definitions of S and L_i can be formalized in $I\Delta_0 + \Omega_1$ and that, for the same reason as in Section 4, $S(1)$ is always defined. However, we do not necessarily have that $S(1) \in K$.

Lemma 6.7. *If $a, b \in K$ and b is an immediate successor of a in $(K_1, <_1)$, then $I\Delta_0 + \Omega_1 \vdash L_a \to \Diamond L_b$.*

Proof. We have $\vdash v(a, b) = \mu x\ \Box_x \neg L_b$ and $\vdash L_a \to v(a, b) = \infty$, whence $\vdash L_a \to \Diamond L_b$ as desired. $\Box$

Lemma 6.8. *If $a \in K$ is a bifurcation point, then $I\Delta_0 + \Omega_1 \vdash L_a \to \Diamond L_{d(a)}$.*

Proof. We have $\vdash v(a, d(a)) = \mu x\ \Box_x \neg L_{d(a)}$. Hence as above $I\Delta_0 + \Omega_1 \vdash L_a \to \Diamond L_{d(a)}$. $\Box$

Lemma 6.9. *If $a <_1 d(a) <_1 b$ and b is an immediate $(<_1)$-successor of $d(a)$, then $I\Delta_0 + \Omega_1 \vdash \Diamond L_{d(a)} \to \Diamond L_b$.*

Proof. Reason in $I\Delta_0 + \Omega_1$. Assume $\Box\neg L_b$. We need to prove $\Box\neg L_{d(a)}$. Let x be such that $\Box_x \neg L_b$. By provable Σ_1^b-completeness, $\Box\Box_x \neg L_b$. Since $\forall u\ \Box(u \in J)$, we have $\Box(\Box_x \neg L_b \wedge x \in J)$. By the small reflection principle $\vdash \forall u\ \Box(L_{d(a)} \to \Diamond_u L_{d(a)})$. So $\Box(L_{d(a)} \to \Diamond_x L_{d(a)} \wedge \Box_x \neg L_b \wedge x \in J)$. By definition, $v(d(a), b) = \mu x \in J\ (\Box_x \neg L_b \wedge \Diamond_x L_{d(a)})$. Thus $\Box(L_{d(a)} \to v(d(a), b) < \infty)$. On the other hand the definition of $L_{d(a)}$ gives us $\Box(L_{d(a)} \to v(d(a), b) = \infty)$. Hence $\Box\neg L_{d(a)}$ as desired. $\Box$

Lemma 6.10. *If $a \in K$, $b \in K_1$, and $a <_1 b$, then $I\Delta_0 + \Omega_1 \vdash L_a \to \Diamond L_b$.*

Proof. By the above lemmas, and by transitivity of 'proves the consistency of'. □

Lemma 6.11. *If $a, b \in K_1$ and $a \not\prec_1 b$, then $I\Delta_0 + \Omega_1 \vdash L_a \rightarrow \Box\neg L_b$.*

Proof. We distinguish the case when $b \preccurlyeq_1 a$ from the case in which b is incomparable with a.

Case 1: Let $b \preccurlyeq_1 a$. From the definitions, $\vdash L_a \rightarrow \Box\neg L_1$. So we can assume $b \neq 1$. Reason in $I\Delta_0 + \Omega_1$. If L_a, then $a = S(1)$, hence by definition, every $b \preccurlyeq_1 a$ with $b \neq 1$ satisfies $w(b) < \infty$. A fortiori $\Box\neg L_b$ as desired.

Case 2: Let b be incomparable with a. It follows that in $(K_1, \prec_1)$ there exists a node of the form $d(i)$ and two immediate $(\prec_1)$-successors u, v of $d(i)$ such that $u \preccurlyeq_1 a$ and $v \preccurlyeq_1 b$. By definition we have $w(u) = \mu x \in J\, (\Box_x \neg L_u \wedge \Diamond_x L_{d(i)})$ and $w(v) = \mu x \in J\, (\Box_x \neg L_v \wedge \Diamond_x L_{d(i)})$. By the properties of the cut J, it follows that $I\Delta_0 + \Omega_1 \vdash w(u) < w(v) \rightarrow \Box(w(u) < w(v))$ and the desired result follows from the fact that $\vdash L_a \rightarrow w(u) < w(v)$ and $\vdash L_b \rightarrow w(v) < w(u)$. □

The next two lemmas follow immediately from the definitions.

Lemma 6.12. *$I\Delta_0 + \Omega_1 \vdash L_i \rightarrow \neg L_j$ for $i \neq j$ in K_1.*

Lemma 6.13. *$I\Delta_0 + \Omega_1 \vdash \Box\neg L_1 \rightarrow \bigvee_{i \in K_1} L_i$.*

Lemma 6.14. *L_1 is consistent with $I\Delta_0 + \Omega_1$.*

Proof. Since for every $i \in K_1$, $I\Delta_0 + \Omega_1 \vdash L_i \rightarrow \Box\neg L_i$, the standard model satisfies $\bigwedge_{i \in K_1} \neg L_i$. On the other hand, by the previous lemma, $\bigwedge_{i \in K_1} \neg L_i$ provably implies $\Diamond L_1$ and the desired result follows. □

We now prove the somewhat surprising:

Lemma 6.15. *If a is a $(\prec_1)$-immediate successor of $d(i)$, then*

$$I\Delta_0 + \Omega_1 \vdash L_{d(i)} \rightarrow \Box\neg L_a.$$

Proof. Recall that $v(d(i), a) = \mu x \in J\, (\Box_x \neg L_a \wedge \Diamond_x L_{d(i)})$. Reason in $I\Delta_0 + \Omega_1$. Assume $L_{d(i)}$. Then there exists x such that $\Box_x \neg L_{d(i)}$. Hence $\Box\Box_x \neg L_{d(i)}$. Reason inside $\Box$. Then $\Box_x \neg L_{d(i)}$ holds. If for a contradiction L_a holds, then $v(d(i), a) < \infty$. Thus there exists y such that $\Box_y \neg L_a$ and $\Diamond_y L_{d(i)}$. It follows that $y < x$. Thus $\Box_x \neg L_a$. Since x is 'external', by the small reflection principle $\neg L_a$ holds. Contradiction. □

Definition 6.16. For $i \in K$, define $k(i)$ as the cardinality of the longest ascending chain in $(K, <)$ whose first element is i. So if i is a leaf, $k(i) = 1$. Extend the map $i \mapsto k(i)$ from K to K_1 by defining $k(d(i)) = k(i) - 1$.

Lemma 6.17. *For every* $u \in K_1$, $I\Delta_0 + \Omega_1 \vdash L_u \rightarrow \Box(\bigvee_{j >_1 u \wedge k(j) < k(u)} L_j)$.

Proof. Since $\vdash L_u \rightarrow \Box\neg L_1$, we have $\vdash L_u \rightarrow \Box(\bigvee_{j \in K_1} L_j)$. So it is enough to show that if j does not satisfy $j >_1 u \wedge k(j) < k(u)$, then $\vdash L_u \rightarrow \Box\neg L_j$. We have already shown that if $\neg(j >_1 u)$, then $\vdash L_u \rightarrow \Box\neg L_j$. On the other hand if $j >_1 u$ and $\neg(k(j) < k(i))$, then u must be of the form $u = d(i)$ and j must be an immediate $(<_i)$-successor of u (hence $k(j) = k(u)$). But then by a previous lemma $\vdash L_u \rightarrow \Box\neg L_j$ as desired. □

Lemma 6.18. *For* $u \in K_1$, $I\Delta_0 + \Omega_1 \vdash L_u \rightarrow \Box^{k(u)}\bot$.
In particular $I\Delta_0 + \Omega_1 \vdash L_1 \rightarrow \Box^m\bot$ *where* m *is the height of* K.

Proof. By induction on $k = k(u)$. The base case is when $k(u) = 1$. Then either u is a leaf, or $u = d(i)$ for some $i \in K$ with $k(i) = 2$. In any case all the nodes $a >_1 u$, if any, are immediate $(<_1)$-successors of u and $\vdash L_u \rightarrow \Box\neg L_a$. But then L_u provably implies $\Box\neg L_j$ for *every* $j \in K_1$, and therefore $\vdash L_u \rightarrow \Box\bot$ as desired. The induction step follows from the previous lemma. □

6.19. $I\Delta_0 + \Omega_1 \vdash L_1 \rightarrow \neg\Box^{m-1}\bot$.

Proof. Clear from the fact that for $a < b$ in K, $\vdash L_a \rightarrow \Diamond L_b$. □

The proof of Theorem 6.2 follows now immediately from all the preceding lemmas.

6.20. If $d(i) <_1 j$ and j is *not* an immediate $(<_1)$-successor of $d(i)$, then we do not know whether $I\Delta_0 + \Omega_1 \vdash L_{d(i)} \rightarrow \Box\neg L_j$ holds, or $I\Delta_0 + \Omega_1 \vdash L_{d(i)} \rightarrow \Diamond L_j$ holds, or neither of them.

References

[1] S. Buss, Bounded Arithmetic (Bibliopolis, edizioni di filosofia e scienze, Napoli, 1986).
[2] A. Berarducci and R. Verbrugge, On the metamathematics of weak theories, ITLI Prepublication Series for Mathematical Logic and Foundations, ML-91-02, University of Amsterdam (1991).
[3] M. Davis, H. Putnam, and J. Robinson, The decision problem for exponential diophantine equations, Ann. of Math. 74 (1961) 425–436.
[4] D. Guaspari and R.M. Solovay, Rosser sentences, Ann. Math. Logic 16 (1979) 81–99.
[5] D. de Jongh, M. Jumelet and F. Montagna, On the proof of Solovay's theorem, Studia Logica 50 (1991).

[6] Yu. Matijasevič, Exponential sets are Diophantine, Dokl. Akad. Nauk SSSR 191 (1970) 279–282. (English translation: Soviet Math. Dokl. (1970) 354–357.)
[7] P. Pudlák, On the length of proofs of finitistic consistency statements in first order theories, in: J.B. Paris, A.J. Wilkie, and G.M. Wilmers, eds., Logic Colloquium '84 (North-Holland, Amsterdam, 1986).
[8] P. Pudlák, Cuts, consistency statements and interpretations, J. Symbolic Logic 50 (2) (1985).
[9] R.M. Solovay, Provability interpretations of modal logic, Israel J. Math. 25 (1976) 287–304.
[10] R.M. Solovay, On interpretability in set theories, Manuscript (1976).
[11] V. Švejdar, Modal analysis of generalized Rosser sentences, J. Symbolic Logic 48 (1983) 986–999.
[12] R. Verbrugge, Docs Solovay's Completeness Theorem Extend to Bounded Arithmetic?, Master's thesis, University of Amsterdam (1988).
[13] R. Verbrugge, Σ-completeness and bounded arithmetic, ITLI Prepublication Series for Mathematical Logic and Foundations, ML-89-05, University of Amsterdam (1989).
[14] A. Visser, On the Σ_1^0-conservativity of Σ_1^0-completeness, Notre Dame J. Formal Logic 32 (1991) 554–561.
[15] A.J. Wilkie and J.B. Paris, On the scheme of induction for bounded arithmetic formulas, Ann. Pure Appl. Logic 35 (1987) 261–302.