



Some Observations on the FGH Theorem

Kurahashi, Taishi

(Citation)

Studia Logica, 111(5):749-778

(Issue Date)

2023-10

(Resource Type)

journal article

(Version)

Version of Record

(Rights)

© The Author 2023

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) a...

(URL)

<https://hdl.handle.net/20.500.14094/0100483303>





TAISHI KURAHASHI 

Some Observations on the FGH Theorem

Abstract. We investigate the Friedman–Goldfarb–Harrington theorem from two perspectives. Firstly, in the frameworks of classical and modal propositional logics, we study the forms of sentences whose existence is guaranteed by the FGH theorem. Secondly, we prove some variations of the FGH theorem with respect to Rosser provability predicates.

Keywords: The FGH theorem, Weak representability of c.e. sets, Incompleteness theorem, Provability logic, Rosser provability predicates.

1. Introduction

The notion of the weak representability of computably enumerable (c.e.) sets plays an important role in a proof of Gödel’s first incompleteness theorem. We say that a set X of natural numbers is *weakly representable* in a theory T if there exists a formula $\varphi_X(v)$ such that for any natural number n , $n \in X$ if and only if $T \vdash \varphi_X(\bar{n})$. If a set X , which is c.e. but not computable, is weakly representable in a c.e. theory T , then it is shown that there exists a natural number n such that $T \not\vdash \varphi_X(\bar{n})$ and $T \not\vdash \neg\varphi_X(\bar{n})$. Therefore, such a theory T is incomplete.

It is easily shown that every c.e. set is weakly representable in every Σ_1 -sound c.e. extension of PA because of Σ_1 -soundness and Σ_1 -completeness. Furthermore, Ehrenfeucht and Feferman [3] proved the weak representability of c.e. sets in every consistent c.e. extension of PA (see also Shepherdson [14]).

THEOREM 1.1. (Ehrenfeucht and Feferman) *Let T be any consistent c.e. extension of PA. Then, every c.e. set of natural numbers is weakly representable in T .*

Ehrenfeucht and Feferman’s theorem is a metamathematical result, but is formalizable and also provable in PA. This fact is called the *FGH theorem* (see Smoryński [15, p.366], Visser [17, Theorem 4.1], [18, Section 3] and Lindström [11, Exercise 2.26 (a)]).

Presented by **Yde Venema**; Received October 28, 2021

THEOREM 1.2. (The FGH theorem) *Let T be any consistent c.e. extension of PA . Then, for any Σ_1 sentence σ , there exists a sentence φ such that*

$$PA + \text{Con}_T \vdash \sigma \leftrightarrow \text{Pr}_T(\ulcorner \varphi \urcorner).$$

Here $\text{Pr}_T(x)$ is a natural Σ_1 provability predicate of T which expresses the provability of T , and Con_T is the Π_1 sentence $\neg \text{Pr}_T(\ulcorner 0 = 1 \urcorner)$ expressing the consistency of T . “FGH” stands for Friedman, Goldfarb and Harrington. Harrington and Friedman pointed out that φ in the statement is found as Π_1 and Σ_1 , respectively. For a history of the FGH theorem, see Visser [18]. The FGH theorem has been used in the literature (for example, in papers by the author, it appears in [8, 10]).

A version of the FGH theorem with a parameter is also proved. That is, by modifying the usual proof of the FGH theorem, it is proved that for any Σ_1 formula $\sigma(v)$ with the only free variable v , there exists a formula $\varphi(v)$ such that $PA + \text{Con}_T$ proves $\forall v (\sigma(v) \leftrightarrow \text{Pr}_T(\ulcorner \varphi(\dot{v}) \urcorner))$. Here $\ulcorner \varphi(\dot{v}) \urcorner$ is a primitive recursive term corresponding to a primitive recursive function calculating the Gödel number of $\varphi(\bar{n})$ from n . If T is consistent, then the theory $PA + \text{Con}_T$ is sound, and hence the weak representability of c.e. sets in T follows from this parameterized version of the FGH theorem.

In the present paper, we analyze the FGH theorem from two perspectives. The first perspective is the “form” of the sentence φ in the statement of the FGH theorem. Let $\text{True}_{\Sigma_1}(v)$ be a partial truth definition for Σ_1 sentences, that is, $\text{True}_{\Sigma_1}(v)$ is a Σ_1 formula such that for any Σ_1 sentence σ , PA proves $\sigma \leftrightarrow \text{True}_{\Sigma_1}(\ulcorner \sigma \urcorner)$ (cf. Lindström [11, Fact 10]). By the parameterized version of the FGH theorem, there exists a formula $\xi(v)$ such that $PA + \text{Con}_T \vdash \forall v (\text{True}_{\Sigma_1}(v) \leftrightarrow \text{Pr}_T(\ulcorner \xi(\dot{v}) \urcorner))$. Then, for any Σ_1 sentence σ , $PA + \text{Con}_T \vdash \sigma \leftrightarrow \text{Pr}_T(\ulcorner \xi(\ulcorner \sigma \urcorner) \urcorner)$. Therefore, φ in the FGH theorem can be taken in the form $\xi(\cdot)$ uniformly regardless of σ .

From the above observation, we consider the following question: What form of the sentence φ in the FGH theorem can we find? In Section 2, we investigate this question in the framework of classical propositional logic. Specifically, we study the following rephrased question: What is a propositional formula A such that φ in the FGH theorem can be taken uniformly in the form A regardless of σ ? In Section 3, we investigate this rephrased question in the framework of modal propositional logic. However, unlike the case of classical propositional logic, the formula A may contain the modal operator \Box . Therefore, the interpretation of \Box in arithmetic should be clearly defined. As is usually done in the study of provability logic, in the present paper, \Box is interpreted by $\text{Pr}_T(x)$. Then, our proofs of theorems in Section 3 are modifications of that of Solovay’s arithmetical completeness theorem [16]

which is one of the most remarkable theorems in the research of provability logic.

The second perspective is the choice of provability predicates. Joosten [7] generalized the FGH theorem by proving similar statements concerning several nonstandard provability predicates such as a formula expressing the provability in T together with all true Σ_{n+2} sentences. In the last section, we study some variations of the FGH theorem with respect to Rosser provability predicates.

Sections 2, 3, and 4 can be read independently. We close the introduction with common preparations for reading these sections. Let \mathcal{L}_A denote the language of first-order arithmetic containing the symbols $0, S, +, \times, \leq$. We do not specify what exactly \mathcal{L}_A is, but it may be assumed to have as many function symbols for primitive recursive functions as necessary. Throughout the present paper, we fix a consistent c.e. extension T of Peano Arithmetic PA in the language \mathcal{L}_A . Let ω denote the set of all natural numbers. For each $n \in \omega$, let \bar{n} denote the numeral $S(S(\cdots S(0) \cdots))$ (n times applications of S) for n . We fix a primitive recursive formula $\text{Proof}_T(x, y)$ naturally expressing that “ y is a T -proof of x ”. Our Σ_1 provability predicate $\text{Pr}_T(x)$ of T is defined by $\exists y \text{Proof}_T(x, y)$, saying that “ x is T -provable”. Then, we may assume that the provability predicate $\text{Pr}_T(x)$ satisfies the following clauses (see Boolos [2]):

- If $T \vdash \varphi$, then $\text{PA} \vdash \text{Pr}_T(\ulcorner \varphi \urcorner)$,
- $\text{PA} \vdash \text{Pr}_T(\ulcorner \varphi \rightarrow \psi \urcorner) \rightarrow (\text{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \text{Pr}_T(\ulcorner \psi \urcorner))$,
- If φ is a Σ_1 sentence, then $\text{PA} \vdash \varphi \rightarrow \text{Pr}_T(\ulcorner \varphi \urcorner)$.

We may also assume that PA verifies that every theorem of T has infinitely many proofs, that is, $\text{PA} \vdash \forall x \forall y (\text{Proof}_T(x, y) \rightarrow \exists z > y \text{Proof}_T(x, z))$.

We introduce witness comparison notation (cf. Lindström [11, Lemma 1.3]).

DEFINITION 1.3. (Witness comparison notation) Suppose that \mathcal{L}_A -formulas φ and ψ are of the forms $\exists x \varphi'(x)$ and $\exists y \psi'(y)$, respectively.

- $\varphi \preceq \psi$ is an abbreviation for $\exists x (\varphi'(x) \wedge \forall y < x \neg \psi'(y))$.
- $\varphi \prec \psi$ is an abbreviation for $\exists x (\varphi'(x) \wedge \forall y \leq x \neg \psi'(y))$.

It is easily verified that PA proves the formulas $\neg(\varphi \preceq \psi \wedge \psi \prec \varphi)$ and $\varphi \vee \psi \rightarrow (\varphi \preceq \psi \vee \psi \prec \varphi)$.

2. On the Form of a Sentence in the FGH Theorem – the Case of Classical Propositional Logic

In this section, we investigate the following question mentioned in the introduction: What is a propositional formula A such that φ in the FGH theorem can be taken uniformly in the form A regardless of σ ?

The language of classical propositional logic consists of countably many propositional variables $p_1, p_2, \dots, q_1, q_2, \dots$, propositional constants \top, \perp , and propositional connectives $\wedge, \vee, \neg, \rightarrow, \leftrightarrow$. For each propositional formula A , let $\models A$ mean that A is a tautology. We say A is unsatisfiable if $\models \neg A$. We say that a propositional formula is *contingent* if it is neither a tautology nor unsatisfiable.

For any propositional formula $A(p_1, \dots, p_n)$ containing only the indicated propositional variables and any \mathcal{L}_A -sentences $\varphi_1, \dots, \varphi_n$, let $A(\varphi_1, \dots, \varphi_n)$ denote the \mathcal{L}_A -sentence obtained by simultaneously replacing all the occurrences of p_i in A by φ_i , for each $i \in \{1, \dots, n\}$.

We introduce the following sets in order to simplify our descriptions.

DEFINITION 2.1. For each Σ_1 sentence σ , let $\text{FGH}_T(\sigma)$ be the set of all \mathcal{L}_A -sentences φ such that $\text{PA} + \text{Con}_T \vdash \sigma \leftrightarrow \text{Pr}_T(\ulcorner \varphi \urcorner)$.

Then, the FGH theorem states that for any Σ_1 sentence σ , the set $\text{FGH}_T(\sigma)$ is non-empty. We show that the set $\text{FGH}(\sigma)$ is closed under the T -provable equivalence.

PROPOSITION 2.2. For any Σ_1 sentence σ and any \mathcal{L}_A -sentences φ and ψ , if $\varphi \in \text{FGH}_T(\sigma)$ and $T \vdash \varphi \leftrightarrow \psi$, then $\psi \in \text{FGH}_T(\sigma)$.

PROOF. Suppose $\varphi \in \text{FGH}_T(\sigma)$ and $T \vdash \varphi \leftrightarrow \psi$. Then, the equivalence $\text{Pr}_T(\ulcorner \varphi \urcorner) \leftrightarrow \text{Pr}_T(\ulcorner \psi \urcorner)$ is provable in PA. Since $\text{PA} + \text{Con}_T \vdash \sigma \leftrightarrow \text{Pr}_T(\ulcorner \varphi \urcorner)$, we obtain $\text{PA} + \text{Con}_T \vdash \sigma \leftrightarrow \text{Pr}_T(\ulcorner \psi \urcorner)$, and hence $\psi \in \text{FGH}_T(\sigma)$. ■

First, we prove the following introductory theorem.

THEOREM 2.3. Let $A(p_1, \dots, p_n)$ be any propositional formula with only the indicated propositional variables. Then, the following are equivalent:

1. $A(p_1, \dots, p_n)$ is contingent.
2. For any Σ_1 sentence σ , there exist \mathcal{L}_A -sentences $\varphi_1, \dots, \varphi_n$ such that $A(\varphi_1, \dots, \varphi_n) \in \text{FGH}_T(\sigma)$.

For each formula A , let A^0 and A^1 be $\neg A$ and A , respectively. Theorem 2.3 follows from the following lemma.

LEMMA 2.4. *Let $A(p_1, \dots, p_n)$ be any propositional formula with only the indicated propositional variables and let r be any propositional variable not contained in A . If $A(p_1, \dots, p_n)$ is contingent, then there exist propositional formulas $B_1(r), \dots, B_n(r)$ such that $\models r \leftrightarrow A(B_1(r), \dots, B_n(r))$.*

PROOF. Suppose that $A(p_1, \dots, p_n)$ is contingent. Let f and g be mappings from $\{1, \dots, n\}$ to $\{0, 1\}$ such that $\models A(\top^{f(1)}, \dots, \top^{f(n)})$ and $\models \neg A(\top^{g(1)}, \dots, \top^{g(n)})$. For each i ($1 \leq i \leq n$), let $B_i(r)$ be the propositional formula $(r \wedge \top^{f(i)}) \vee (\neg r \wedge \top^{g(i)})$. Then, $\models r \rightarrow (B_i(r) \leftrightarrow \top^{f(i)})$ and $\models \neg r \rightarrow (B_i(r) \leftrightarrow \top^{g(i)})$ hold. Thus, we have

$$\models r \rightarrow (A(B_1(r), \dots, B_n(r)) \leftrightarrow A(\top^{f(1)}, \dots, \top^{f(n)}))$$

and

$$\models \neg r \rightarrow (A(B_1(r), \dots, B_n(r)) \leftrightarrow A(\top^{g(1)}, \dots, \top^{g(n)})).$$

Since $\models A(\top^{f(1)}, \dots, \top^{f(n)})$ and $\models \neg A(\top^{g(1)}, \dots, \top^{g(n)})$, we obtain

$$\models r \rightarrow A(B_1(r), \dots, B_n(r)) \text{ and } \models \neg r \rightarrow \neg A(B_1(r), \dots, B_n(r)).$$

■

Proof of Theorem 2.3. ($1 \Rightarrow 2$): Suppose that $A(p_1, \dots, p_n)$ is contingent and let σ be any Σ_1 sentence. Then, by Lemma 2.4, there exist propositional formulas $B_1(r), \dots, B_n(r)$ such that

$$\models r \leftrightarrow A(B_1(r), \dots, B_n(r)). \quad (\text{I})$$

By the FGH theorem, there exists an \mathcal{L}_A -sentence $\varphi \in \text{FGH}_T(\sigma)$. For each i ($1 \leq i \leq n$), let φ_i be the \mathcal{L}_A -sentence $B_i(\varphi)$. Then, by the equivalence (I), we obtain

$$\text{PA} \vdash \varphi \leftrightarrow A(\varphi_1, \dots, \varphi_n).$$

By Proposition 2.2, we conclude $A(\varphi_1, \dots, \varphi_n) \in \text{FGH}_T(\sigma)$ by Proposition 2.2.

($2 \Rightarrow 1$): Suppose that $A(p_1, \dots, p_n)$ is not contingent. Then, for any \mathcal{L}_A -sentences $\varphi_1, \dots, \varphi_n$, either $A(\varphi_1, \dots, \varphi_n)$ or $\neg A(\varphi_1, \dots, \varphi_n)$ is provable in PA. Let σ be any Σ_1 sentence independent of $\text{PA} + \text{Con}_T$. Then, $A(\varphi_1, \dots, \varphi_n)$ is not in $\text{FGH}_T(\sigma)$ for all \mathcal{L}_A -sentences $\varphi_1, \dots, \varphi_n$. ■

As mentioned in the introduction, the FGH theorem and the weak representability theorem are related to each other. In particular, we can also prove the following theorem by a similar proof.

THEOREM 2.5. *Let $A(p_1, \dots, p_n)$ be any propositional formula with only the indicated propositional variables. Then, the following are equivalent:*

1. $A(p_1, \dots, p_n)$ is contingent.
2. For any c.e. set X , there exist \mathcal{L}_A -formulas $\varphi_1(v), \dots, \varphi_n(v)$ such that $A(\varphi_1(v), \dots, \varphi_n(v))$ weakly represents X in T .

Theorem 2.3 will be extended to the framework of modal propositional logic in the next section. In this section, we further improve Theorem 2.3 in the framework of classical propositional logic. For any propositional formula $A(p_1, \dots, p_n, q_1, \dots, q_m)$ with the only indicated propositional variables, let $\text{FGH}_T[A; q_1, \dots, q_m]$ denote the metamathematical statement “for any \mathcal{L}_A -sentences ψ_1, \dots, ψ_m and for any Σ_1 sentence σ , there exist \mathcal{L}_A -sentences $\varphi_1, \dots, \varphi_n$ such that $A(\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m) \in \text{FGH}_T(\sigma)$ ”, and we provide a necessary and sufficient condition for $\text{FGH}_T[A; q_1, \dots, q_m]$. From this, we obtain more detailed information about the elements of $\text{FGH}(\sigma)$ and the first incompleteness theorem (see Corollary 2.8).

Let \mathcal{F}_m denote the set of all functions $f : \{1, \dots, m\} \rightarrow \{0, 1\}$. We prove the following theorem which is one of main theorems of the present paper.

THEOREM 2.6. *For any propositional formula $A(p_1, \dots, p_n, q_1, \dots, q_m)$, the following are equivalent:*

1. For all $f \in \mathcal{F}_m$, $A(p_1, \dots, p_n, \top^{f(1)}, \dots, \top^{f(m)})$ are contingent.
2. $\text{FGH}_T[A; q_1, \dots, q_m]$ holds.

Theorem 2.6 also follows from the following lemma that is an improvement of Lemma 2.4. For the sake of simplicity, we sometimes abbreviate the tuples p_1, \dots, p_n and q_1, \dots, q_m as \vec{p} and \vec{q} , respectively.

LEMMA 2.7. *Let $A(\vec{p}, \vec{q})$ be any propositional formula with only the indicated propositional variables and let r be any propositional variable not contained in A . If $A(\vec{p}, \top^{f(1)}, \dots, \top^{f(m)})$ is contingent for all $f \in \mathcal{F}_m$, then there exist propositional formulas $B_1(r, \vec{q}), \dots, B_n(r, \vec{q})$ such that*

$$\models r \leftrightarrow A(B_1(r, \vec{q}), \dots, B_n(r, \vec{q}), \vec{q}).$$

PROOF. For each $f \in \mathcal{F}_m$, let $Q^f(\vec{q})$ denote the formula $q_1^{f(1)} \wedge \dots \wedge q_m^{f(m)}$. Notice that for each distinct elements $f, g \in \mathcal{F}_m$, $\models \neg(Q^f(\vec{q}) \wedge Q^g(\vec{q}))$. Also $\models \bigvee_{f \in \mathcal{F}_m} Q^f(\vec{q})$.

Suppose that $A(\vec{p}, \top^{f(1)}, \dots, \top^{f(m)})$ is contingent for all $f \in \mathcal{F}_m$. For each $g \in \mathcal{F}_m$, there exist propositional formulas $C_1^g(r), \dots, C_n^g(r)$ such that

$$\models r \leftrightarrow A(C_1^g(r), \dots, C_n^g(r), \top^{g(1)}, \dots, \top^{g(m)}) \quad (\text{II})$$

by Lemma 2.4. For each i ($1 \leq i \leq n$), let $B_i(r, \vec{q})$ be the propositional formula $\bigvee_{f \in \mathcal{F}_m} (C_i^f(r) \wedge Q^f(\vec{q}))$. Then, $\models Q^g(\vec{q}) \rightarrow (B_i(r, \vec{q}) \leftrightarrow C_i^g(r))$ and

$\models Q^g(\vec{q}) \rightarrow (q_i \leftrightarrow \top^{g(i)})$. By the equivalence (II), we obtain

$$\models Q^g(\vec{q}) \rightarrow (r \leftrightarrow A(B_1(r, \vec{q}), \dots, B_n(r, \vec{q}), \vec{q})).$$

Then,

$$\models \bigvee_{f \in \mathcal{F}_m} Q^f(\vec{q}) \rightarrow (r \leftrightarrow A(B_1(r, \vec{q}), \dots, B_n(r, \vec{q}), \vec{q})).$$

Since $\models \bigvee_{f \in \mathcal{F}_m} Q^f(\vec{q})$, we conclude

$$\models r \leftrightarrow A(B_1(r, \vec{q}), \dots, B_n(r, \vec{q}), \vec{q}).$$

■

Proof of Theorem 2.6. ($1 \Rightarrow 2$): Suppose that $A(\vec{p}, \top^{f(1)}, \dots, \top^{f(m)})$ is contingent for all $f \in \mathcal{F}_m$. Let ψ_1, \dots, ψ_m be any \mathcal{L}_A -sentences and let σ be any Σ_1 sentence. By Lemma 2.7, there exist propositional formulas $B_1(r, \vec{q}), \dots, B_n(r, \vec{q})$ such that

$$\models r \leftrightarrow A(B_1(r, \vec{q}), \dots, B_n(r, \vec{q}), \vec{q}). \quad (\text{III})$$

By the FGH theorem, there exists an \mathcal{L}_A -sentence $\varphi \in \text{FGH}_T(\sigma)$. For each i ($1 \leq i \leq n$), let φ_i be the \mathcal{L}_A -sentence $B_i(\varphi, \psi_1, \dots, \psi_m)$. Then, by the equivalence (III),

$$\text{PA} \vdash \varphi \leftrightarrow A(\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m).$$

By Proposition 2.2, we have $A(\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m) \in \text{FGH}_T(\sigma)$. Therefore, we conclude that $\text{FGH}_T[A; \vec{q}]$ holds.

($2 \Rightarrow 1$): Suppose that $A(p_1, \dots, p_n, \top^{f(1)}, \dots, \top^{f(m)})$ is either a tautology or unsatisfiable for some $f \in \mathcal{F}_m$. For $i \in \{1, \dots, m\}$, let ψ_i be the \mathcal{L}_A -sentence $0 = 0^{f(i)}$. Then, for any \mathcal{L}_A -sentences $\varphi_1, \dots, \varphi_n$, the \mathcal{L}_A -sentence $A(\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m)$ is either provable or refutable in PA. Let σ be any Σ_1 sentence independent of $\text{PA} + \text{Con}_T$. Then, $A(\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m) \notin \text{FGH}_T(\sigma)$. Therefore $\text{FGH}_T[A; q_1, \dots, q_m]$ does not hold. ■

For example, let $A(p, q)$ be the propositional formula $p \leftrightarrow q$. Since both $p \leftrightarrow \top$ and $p \leftrightarrow \perp$ are contingent, $\text{FGH}_T[A; q]$ holds by Theorem 2.6. That is, for any \mathcal{L}_A -sentence ψ and any Σ_1 sentence σ , there exists an \mathcal{L}_A -sentence φ such that $\varphi \leftrightarrow \psi \in \text{FGH}_T(\sigma)$. Another interesting corollary to Theorem 2.6 is the following version of the first incompleteness theorem.

COROLLARY 2.8. *Let $A(p_1, \dots, p_n, q_1, \dots, q_m)$ be any propositional formula. If $A(p_1, \dots, p_n, \top^{f(1)}, \dots, \top^{f(m)})$ are contingent for all $f \in \mathcal{F}_m$, then for any \mathcal{L}_A -sentences ψ_1, \dots, ψ_m , there exist \mathcal{L}_A -sentences $\varphi_1, \dots, \varphi_n$ such that $A(\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m)$ is independent of T .*

PROOF. Let σ be a Σ_1 sentence independent of $\text{PA} + \text{Con}_T$ and ψ_1, \dots, ψ_m be any \mathcal{L}_A -sentences. By Theorem 2.6, there exist \mathcal{L}_A -sentences $\varphi_1, \dots, \varphi_n$ such that

$$\text{PA} + \text{Con}_T \vdash \sigma \leftrightarrow \text{Pr}_T(\ulcorner A(\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m) \urcorner).$$

Then, it is easy to show that $A(\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m)$ is independent of T . ■

Notice that we can also prove the parameterized version and the weak representability version of Theorem 2.6.

3. On the Form of a Sentence in the FGH Theorem – the Case of Modal Propositional Logic

In this section, we extend Theorem 2.3 to the framework of modal propositional logic. This section consists of three subsections. First, we give some preparations which are needed for our arguments. In the second subsection, we prove our modal version of Theorem 2.3. Finally, we also investigate our modal version for \mathcal{L}_A -theories with finite heights.

Here we define the height of theories. The sequence $\langle \text{Con}_T^n \rangle_{n \in \omega}$ of \mathcal{L}_A -sentences is recursively defined as follows: Con_T^0 is $0 = 0$; and Con_T^{n+1} is $\neg \text{Pr}_T(\ulcorner \neg \text{Con}_T^n \urcorner)$. Notice that Con_T^1 is PA -provably equivalent to Con_T . If there exists a natural number $k \geq 1$ such that $T \vdash \neg \text{Con}_T^k$, then the least such a number k is called the *height* of T . If not, we say that the height of T is ∞ .

3.1. Preparations

The language of modal propositional logic is that of classical propositional logic equipped with the unary modal operation \Box . For each modal formula A , let $\text{Sub}(A)$ be the set of all subformulas of A . Let $\mathfrak{c}(A)$ be the cardinality of the set $\{B \in \text{Sub}(A) \mid B \text{ is of the form } \Box C\}$. The modal formula $\Box^n A$ is recursively defined as follows: $\Box^0 A$ is A ; and $\Box^{n+1} A$ is $\Box \Box^n A$. The formula $\Diamond^n A$ is an abbreviation for $\neg \Box^n \neg A$.

The base logic of our investigations in this section is the Gödel–Löb logic **GL** which is known as the logic of provability. The axioms of **GL** are as follows:

1. All propositional tautologies in the language of modal propositional logic,
2. $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$,

3. $\Box(\Box p \rightarrow p) \rightarrow \Box p$.

The inference rules of **GL** are modus ponens, necessitation, and uniform substitution.

A **GL-frame** is a tuple $\langle W, \sqsubset, r \rangle$ where W is a nonempty finite set, \sqsubset is a transitive irreflexive binary relation on W and r is an element of W with $r \sqsubset x$ for all $x \in W \setminus \{r\}$. Such an element r is called the *root* of the frame. A **GL-model** is a tuple $M = \langle W, \sqsubset, r, \Vdash \rangle$ where $\langle W, \sqsubset, r \rangle$ is a **GL-frame**, and \Vdash is a binary relation between W and the set of all modal formulas satisfying the usual conditions for satisfaction and the following condition: $a \Vdash \Box A$ if and only if for all $b \in W$, $b \Vdash A$ if $a \sqsubset b$. It is known that **GL** is sound and complete with respect to the class of all **GL-frames** (cf. Segerberg [12]). Moreover, the proof of the Kripke completeness of **GL** given in the textbook [2] by Boolos shows that the following theorem holds.

THEOREM 3.1. *For any modal formula A , if $\mathbf{GL} \not\vdash A$, then there exists a **GL-model** $\langle W, \sqsubset, r, \Vdash \rangle$ such that $r \Vdash \Box^{c(A)+1} \perp$ and $r \not\vdash A$.*

For each set Γ of modal formulas, let $\mathbf{GL} + \Gamma$ denote the logic whose axioms are all theorems of **GL** and all elements of Γ , and whose inference rules are modus ponens and uniform substitution. We identify each axiomatic system of modal propositional logic with the set of all its theorems. We introduce the following two extensions of **GL** which are studied in the context of the classification of propositional provability logics (cf. Artemov and Beklemishev [1]).

- $\mathbf{GL}_\omega := \mathbf{GL} + \{\Diamond^n \top \mid n \in \omega\}$.
- $\mathbf{S} := \mathbf{GL} + \{\Box p \rightarrow p\}$.

Notice $\mathbf{GL} \subseteq \mathbf{GL}_\omega \subseteq \mathbf{S}$. To connect these logics with arithmetic, we introduce the notion of arithmetical interpretation.

DEFINITION 3.2. (Arithmetical interpretations) A mapping from the set of all propositional variables to a set of \mathcal{L}_A -sentences is called an *arithmetical interpretation*. Each arithmetical interpretation f is uniquely extended to the mapping f_T from the set of all modal formulas to a set of \mathcal{L}_A -sentences by the following clauses:

1. $f_T(\perp)$ is $0 = 1$,
2. $f_T(\neg A)$ is $\neg f_T(A)$,
3. $f_T(A \circ B)$ is $f_T(A) \circ f_T(B)$ for $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$,
4. $f_T(\Box A)$ is $\text{Pr}_T(\ulcorner f_T(A) \urcorner)$.

The logics **GL**, **GL_ω** and **S** are sound with respect to arithmetical interpretations. Let \mathbb{N} denote the standard model of arithmetic.

FACT 3.3. (Arithmetical soundness (cf. Artemov and Beklemishev [1])) *Let A be any modal formula.*

1. *If $\mathbf{GL} \vdash A$, then $PA \vdash f_T(A)$ for any arithmetical interpretation f .*
2. *If the height of T is ∞ and $\mathbf{GL}_\omega \vdash A$, then $\mathbb{N} \models f_T(A)$ for any arithmetical interpretation f .*
3. *If T is sound and $\mathbf{S} \vdash A$, then $\mathbb{N} \models f_T(A)$ for any arithmetical interpretation f .*

We show some interrelationships between these logics. For each modal formula A , let $\text{Rfn}(A)$ denote the set $\{\Box B \rightarrow B \mid \Box B \in \text{Sub}(A)\}$. Notice that the cardinality of the set $\text{Rfn}(A)$ is exactly $\mathfrak{c}(A)$.

FACT 3.4. (Solovay [16]) *For any modal formula A , the following are equivalent:*

1. $\mathbf{S} \vdash A$.
2. $\mathbf{GL} \vdash \bigwedge \text{Rfn}(A) \rightarrow A$.

The following fact is a kind of folklore which is proved through arithmetical interpretations.

FACT 3.5. *For any modal formula A , the following are equivalent:*

1. $\mathbf{GL} \vdash A$.
2. $\mathbf{GL}_\omega \vdash \Box A$.
3. $\mathbf{S} \vdash \Box A$.

LEMMA 3.6. *For any modal formula A , the following are equivalent:*

1. $\mathbf{GL} \vdash \bigwedge \text{Rfn}(\Box A) \rightarrow \neg \Box A$.
2. $\mathbf{GL} \vdash \Diamond^{\mathfrak{c}(A)+1} \top \rightarrow \neg \Box A$.
3. $\mathbf{GL}_\omega \vdash \neg \Box A$.
4. $\mathbf{S} \vdash \neg \Box A$.

PROOF. (1 \Rightarrow 2): Suppose $\mathbf{GL} \vdash \bigwedge \text{Rfn}(\Box A) \rightarrow \neg \Box A$. It is known that **GL** proves $\Diamond^{\mathfrak{c}(\Box A)} \top \rightarrow \bigwedge \text{Rfn}(\Box A) \vee \Diamond \bigwedge \text{Rfn}(\Box A)$ (cf. [1, Lemma 26]). Then, $\mathbf{GL} \vdash \Diamond^{\mathfrak{c}(\Box A)} \top \rightarrow \neg \Box A \vee \Diamond \neg \Box A$. Since $\mathbf{GL} \vdash \Diamond \neg \Box A \rightarrow \neg \Box A$, we have $\mathbf{GL} \vdash \Diamond^{\mathfrak{c}(\Box A)} \top \rightarrow \neg \Box A$. Hence, $\mathbf{GL} \vdash \Diamond^{\mathfrak{c}(A)+1} \top \rightarrow \neg \Box A$ because $\mathfrak{c}(\Box A) = \mathfrak{c}(A) + 1$.

(2 \Rightarrow 3) and (3 \Rightarrow 4): Obvious.

(4 \Rightarrow 1): By Fact 3.4 and $\text{Rfn}(\Box A) = \text{Rfn}(\neg\Box A)$. ■

The conditions “ $\mathbf{GL} \vdash A$ ” and “ $\mathbf{GL}_\omega \vdash \neg\Box A$ ” are also characterized by provability in other extensions of \mathbf{GL} .

DEFINITION 3.7. Let F_s be the modal formula $\Box^{s+1}\perp \rightarrow \Box^s\perp$.

LEMMA 3.8. For any modal formula A and any natural number s with $s > \mathfrak{c}(A)$, the following are equivalent:

1. $\mathbf{GL} \vdash A$.
2. $\mathbf{GL} + \{\neg F_s\} \vdash \Box A$.

PROOF. (1 \Rightarrow 2): Obvious.

(2 \Rightarrow 1): If $\mathbf{GL} \not\vdash A$, then there exists a \mathbf{GL} -model $\langle W, \Box, r, \Vdash \rangle$ such that $r \Vdash \Box^{\mathfrak{c}(A)+1}\perp \wedge \neg A$ by Theorem 3.1. Let $\langle W', \Box', r', \Vdash' \rangle$ be a \mathbf{GL} -model obtained from $\langle W, \Box, r, \Vdash \rangle$ by adding a chain of new elements below r so that $r' \Vdash' \Box^{s+1}\perp \wedge \Diamond^s\top$. Such a model exists because $s > \mathfrak{c}(A)$. Since $r' \Vdash' \neg\Box A$, we obtain $\mathbf{GL} + \{\neg F_s\} \not\vdash \Box A$. ■

LEMMA 3.9. For any modal formula A and any natural number s with $s > \mathfrak{c}(A)$, the following are equivalent:

1. $\mathbf{GL}_\omega \vdash \neg\Box A$.
2. $\mathbf{GL} + \{\neg F_s\} \vdash \neg\Box A$.

PROOF. (1 \Rightarrow 2): If $\mathbf{GL}_\omega \vdash \neg\Box A$, then by Lemma 3.6, $\Diamond^{\mathfrak{c}(A)+1}\top \rightarrow \neg\Box A$ is proved in \mathbf{GL} , and thus $\mathbf{GL} + \{\neg F_s\} \vdash \neg\Box A$ because $s \geq \mathfrak{c}(A) + 1$.

(2 \Rightarrow 1): Suppose $\mathbf{GL} + \{\neg F_s\} \vdash \neg\Box A$, then $\mathbf{GL} \vdash \Box A \rightarrow F_s$, that is, $\mathbf{GL} \vdash \Box A \rightarrow (\Box^{s+1}\perp \rightarrow \Box^s\perp)$. Then, $\mathbf{GL} \vdash \Box\Box A \rightarrow \Box(\Box^{s+1}\perp \rightarrow \Box^s\perp)$. By Löb's principle, we get $\mathbf{GL} \vdash \Box\Box A \rightarrow \Box^{s+1}\perp$. Then, $\mathbf{GL} \vdash \Box A \rightarrow \Box^{s+1}\perp$, and hence $\mathbf{GL} \vdash \Diamond^{s+1}\top \rightarrow \neg\Box A$. We obtain $\mathbf{GL}_\omega \vdash \neg\Box A$. ■

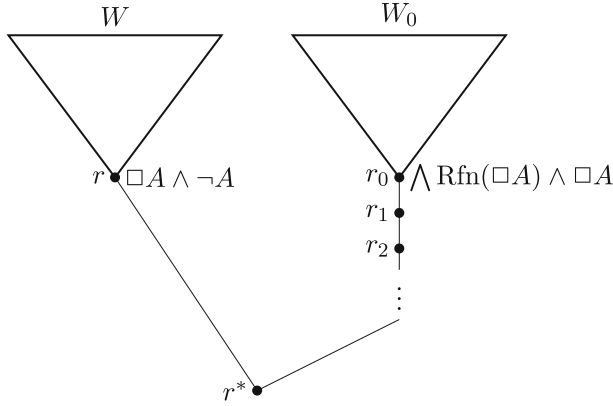
Here we introduce a notion that will be central in this section.

DEFINITION 3.10. A modal formula A is said to be *nontrifling* if $\mathbf{GL}_\omega \not\vdash \Box\Box A \rightarrow \Box A$.

Then, this notion is characterized in many ways as follows.

PROPOSITION 3.11. For any modal formula A and any number $s > \mathfrak{c}(A)$, the following are equivalent:

1. A is nontrifling.
2. $\mathbf{GL}_\omega \not\vdash \Box A$ and $\mathbf{GL}_\omega \not\vdash \neg\Box A$.

Figure 1. The **GL**-model M^*

3. $\mathbf{S} \not\models \Box A$ and $\mathbf{S} \not\models \neg \Box A$.

4. $\mathbf{GL} + \{\neg F_s\} \not\models \Box A$ and $\mathbf{GL} + \{\neg F_s\} \not\models \neg \Box A$.

5. $\mathbf{GL} \not\models A$ and $\mathbf{GL} \not\models \bigwedge \text{Rfn}(\Box A) \rightarrow \neg \Box A$.

PROOF. ($1 \Rightarrow 2$): We prove the contrapositive. If $\mathbf{GL}_\omega \vdash \Box A$, then $\mathbf{GL}_\omega \vdash \Box \Box A \rightarrow \Box A$ is obvious. Suppose $\mathbf{GL}_\omega \vdash \neg \Box A$. By Lemma 3.6, we have $\mathbf{GL} \vdash \Diamond^{c(A)+1} \top \rightarrow \neg \Box A$. Then, $\mathbf{GL} \vdash \Diamond^{c(A)+2} \top \rightarrow \neg \Box \Box A$. We obtain $\mathbf{GL}_\omega \vdash \neg \Box \Box A$, and hence $\mathbf{GL}_\omega \vdash \Box \Box A \rightarrow \Box A$.

($2 \Leftrightarrow 3 \Leftrightarrow 4 \Leftrightarrow 5$): By Fact 3.5 and Lemmas 3.6, 3.8, and 3.9.

($5 \Rightarrow 1$): Suppose $\mathbf{GL} \not\models A$ and $\mathbf{GL} \not\models \bigwedge \text{Rfn}(\Box A) \rightarrow \neg \Box A$. Then, $\mathbf{GL} \not\models \Box A \rightarrow A$. By Theorem 3.1, there exist **GL**-models $M = \langle W, \sqsubset, r, \Vdash \rangle$ and $M_0 = \langle W_0, \sqsubset_0, r_0, \Vdash_0 \rangle$ such that $r \Vdash \Box A \wedge \neg A$ and $r_0 \Vdash_0 \bigwedge \text{Rfn}(\Box A) \wedge \Box A$. We merge these two models M and M_0 into one model M^* , which is illustrated in Fig 1. We give the precise definition of the **GL**-model $M^* = \langle W^*, \sqsubset^*, r^*, \Vdash^* \rangle$ as follows:

- $W^* = W \cup W_0 \cup \{r_i \mid i \geq 1\} \cup \{r^*\}$,

- \sqsubset^* is the transitive closure of

$$\sqsubset \cup \sqsubset_0 \cup \{(r_i, r_j) \mid i > j\} \cup \{(r^*, r_i) \mid i \in \omega\} \cup \{(r^*, r)\},$$

- for $a \in W$, $a \Vdash^* p$ if and only if $a \Vdash p$,

for $a \in W_0$, $a \Vdash^* p$ if and only if $a \Vdash_0 p$, and

for $a \in \{r_i \mid i \geq 1\} \cup \{r^*\}$, $a \Vdash^* p$ if and only if $r_0 \Vdash_0 p$.

It is easy to show that $r \Vdash^* \Box A \wedge \neg A$ and $r_0 \Vdash^* \bigwedge \text{Rfn}(\Box A) \wedge \Box A$. Since $r^* \Vdash^* \Diamond^n \top$ for all $n \in \omega$, every theorem of \mathbf{GL}_ω is true in r^* . Thus, it suffices

to show that $r^* \not\models^* \Box\Box A \rightarrow \Box A$. For this purpose, we prove the following lemma.

LEMMA 3.12. *For any $i \in \omega$ and any subformula B of $\Box A$, $r_i \models^* B$ if and only if $r_0 \models^* B$.*

PROOF. We prove the lemma by induction on i . For $i = 0$, the lemma is trivial. Suppose that the lemma holds for i , and we prove the case of $i + 1$ by induction on the construction of the subformula B of $\Box A$. The case that B is a propositional variable is obvious from the definition of \models^* . The cases for propositional connectives are easily proved by the induction hypothesis. We only give a proof of the case that B is of the form $\Box C$, that is, we prove $r_{i+1} \models^* \Box C$ if and only if $r_0 \models^* \Box C$.

(\Rightarrow): Suppose $r_{i+1} \models^* \Box C$, then $r_{i+1} \models^* \Box\Box C$ because \sqsubset^* is transitive. Since $r_{i+1} \sqsubset^* r_0$, we obtain $r_0 \models^* \Box C$.

(\Leftarrow): Suppose $r_{i+1} \not\models^* \Box C$. Then, $a \not\models^* C$ for some $a \in W^*$ with $r_{i+1} \sqsubset^* a$. We distinguish the following three cases:

- $r_0 \sqsubset^* a$.

Then, $r_0 \not\models^* \Box C$ trivially holds.

- $a = r_0$.

We have $r_0 \models^* \Box C \rightarrow C$ because $r_0 \models^* \text{Rfn}(\Box A)$. Thus, $r_0 \not\models^* \Box C$.

- $a = r_j$ for $0 < j < i + 1$.

By the induction hypothesis for j , we have $r_0 \not\models^* C$. Then, $r_j \not\models^* \Box C$. By the induction hypothesis for j again, we obtain $r_0 \not\models^* \Box C$. ■

We are ready to prove $r^* \not\models^* \Box\Box A \rightarrow \Box A$. Let $a \in W^*$ be any element such that $r^* \sqsubset^* a$. If $a \in W$, then $a \models^* \Box A$ because $r \models^* \Box A$. If $a \in W_0$, then $a \models^* \Box A$ because $r_0 \models^* \Box A$. If $a \in \{r_i \mid i \geq 1\}$, then $a \models^* \Box A$ by Lemma 3.12. Therefore, we obtain $r^* \models^* \Box\Box A$. Since $r^* \sqsubset^* r$ and $r \not\models^* A$, we get $r^* \not\models^* \Box A$. Thus, we conclude $r^* \not\models^* \Box\Box A \rightarrow \Box A$, and hence A is nontrifling. ■

If a formula A does not contain \Box , then the following proposition holds.

PROPOSITION 3.13. *A propositional formula A is contingent if and only if A is nontrifling.*

PROOF. (\Rightarrow): Let v_0 be a truth assignment in classical propositional logic such that $v_0(A)$ is false. Let $\langle W_0, \sqsubset_0, r_0, \models_0 \rangle$ be a **GL**-model satisfying $W_0 = \{r_0\}$, $\sqsubset_0 = \emptyset$, and $r_0 \models_0 p$ if and only if $v_0(p)$ is true. Then $r_0 \not\models_0 A$, and hence **GL** $\not\models A$.

Let v_1 be a truth assignment such that $v_1(A)$ is true. Let $\langle W_1, \sqsubset_1, r_1, \Vdash_1 \rangle$ be a **GL**-model satisfying $W_1 = \{r_1, a\}$, $r_1 \sqsubset_1 a$, and for each $w \in W_1$, $w \Vdash_1 p$ if and only if $v_1(p)$ is true. Then, $r_1 \Vdash_1 \Diamond \top \wedge \Box A$. Since $c(A) = 0$, we have $\mathbf{GL} \not\vdash \Diamond^{c(A)+1} \top \rightarrow \neg \Box A$. By Lemma 3.6, $\mathbf{GL}_\omega \not\vdash \neg \Box A$. By Proposition 3.11, A is nontrifling.

(\Leftarrow): If A is a tautology, then $\mathbf{GL} \vdash A$. If A is unsatisfiable, then $\mathbf{GL} \vdash \neg A$, and $\mathbf{GL} \vdash \Diamond \top \rightarrow \neg \Box A$. Hence, $\mathbf{GL}_\omega \vdash \neg \Box A$. In either case, A is not nontrifling by Proposition 3.11. ■

3.2. A modal version of the FGH Theorem

We are ready to prove an extension of Theorem 2.3.

THEOREM 3.14. *For any modal formula A , if A is nontrifling, then for any Σ_1 sentence σ , there exists an arithmetical interpretation f such that*

- (a) $PA \vdash \sigma \rightarrow f_T(\Box A)$, and
- (b) $PA + \text{Con}_T^{c(A)+1} \vdash f_T(\Box A) \rightarrow \sigma$.

PROOF. Suppose that A is nontrifling and let σ be any Σ_1 sentence. We may assume that σ is of the form $\exists x \delta(x)$ for some Δ_0 formula $\delta(x)$. By Proposition 3.11, we have $\mathbf{GL} \not\vdash A$ and $\mathbf{GL} \not\vdash \bigwedge \text{Rfn}(\Box A) \rightarrow \neg \Box A$. By Theorem 3.1, there exist **GL**-models $M_0 = \langle W_0, \sqsubset_0, r_0, \Vdash_0 \rangle$ and $M_1 = \langle W_1, \sqsubset_1, r_1, \Vdash_1 \rangle$ such that $r_0 \Vdash_0 \Box^{c(A)+1} \perp \wedge \neg A$ and $r_1 \Vdash_1 \bigwedge \text{Rfn}(\Box A) \wedge \Box A$.

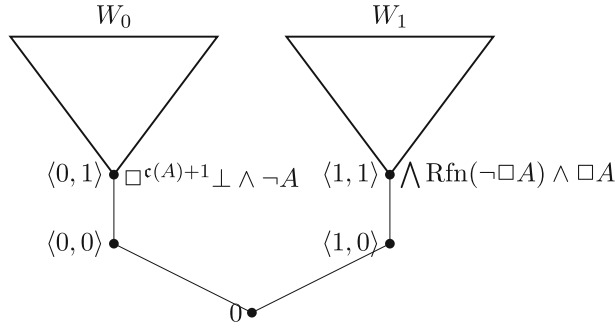
Let $\langle \cdot, \cdot \rangle$ be a natural injective mapping from ω^2 to ω . Also let π_0 and π_1 be natural mappings such that $\pi_0(\langle i, j \rangle) = i$ and $\pi_1(\langle i, j \rangle) = j$. We may assume that 0 is not in the range of the mapping $\langle \cdot, \cdot \rangle$ and is not in the domain of the mappings π_0 and π_1 . Let n_0 and n_1 be the cardinalities of W_0 and W_1 , respectively. Without loss of generality, we may assume

- $W_0 = \{\langle 0, 1 \rangle, \langle 0, 2 \rangle, \dots, \langle 0, n_0 \rangle\}$, $r_0 = \langle 0, 1 \rangle$, and
- $W_1 = \{\langle 1, 1 \rangle, \langle 1, 2 \rangle, \dots, \langle 1, n_1 \rangle\}$, $r_1 = \langle 1, 1 \rangle$.

We merge the two models M_0 and M_1 into one **GL**-model M , which is visualized in Fig. 2. The definition of $M = \langle W', \sqsubset', 0, \Vdash' \rangle$ is as follows:

- $W' = W_0 \cup W_1 \cup \{0, \langle 0, 0 \rangle, \langle 1, 0 \rangle\}$,
- $\sqsubset' = \sqsubset_0 \cup \sqsubset_1 \cup \{ \langle 0, a \rangle \mid a \in W' \setminus \{0\} \} \cup \{ \langle \langle i, 0 \rangle, a \rangle \mid i \in \{0, 1\}, a \in W_i \}$,
- For $\langle i, j \rangle \in W_i$, $\langle i, j \rangle \Vdash' p$ if and only if $\langle i, j \rangle \Vdash_i p$.

Also $0 \Vdash' p$, and $\langle i, 0 \rangle \Vdash' p$ if and only if $\langle i, 1 \rangle \Vdash_i p$.

Figure 2. The **GL**-model M'

Then, it is easy to show that $\langle 0, 1 \rangle \Vdash' \Box^{c(A)+1} \perp \wedge \neg A$ and $\langle 1, 1 \rangle \Vdash' \bigwedge \text{Rfn}(\neg \Box A) \wedge \Box A$. As in the usual proof of Solovay's arithmetical completeness theorem, we recursively define a primitive recursive function $h : \omega \rightarrow W'$ by referring to T -proofs. In the definition of h , we use the formula $\lambda(x) \equiv \exists y \forall z \geq y (h(z) = x)$ and the sentence $f_T(A)$ for the arithmetical interpretation f defined by $f(p) \equiv \bigvee_{a \in W', a \Vdash' p} \lambda(\bar{a})$. This is done by the aid of the Fixed Point Lemma or the recursion theorem as in the proof of Solovay's theorem because such $\lambda(x)$ and $f_T(A)$ are effectively computable from h (see Boolos [2]).

Here we give the definition of the function h . Let $h(0) := 0$. Suppose that the value of $h(x)$ has already been defined. We define the value of $h(x+1)$. If $\forall y \leq x \neg \delta(y)$ holds and there exists no T -proof of $f_T(A)$ smaller than or equal to x , let $h(x+1) := 0$. That is, the output of h remains 0 unless the smallest witness of $\delta(x)$ or the smallest T -proof of $f_T(A)$ appears.

After such a witness appears, h changes its value. At the first stage when the smallest witness x of $\delta(x)$ or the smallest T -proof x of $f_T(A)$ appears, we distinguish the following two cases:

- Case 1: $\forall y < x \neg \delta(y)$ holds and x is the smallest T -proof of $f_T(A)$.

Let $h(x+1) := \langle 0, 0 \rangle$.

- Case 2: $\forall y < x \neg \delta(y)$ and $\delta(x)$ hold and there is no proof of $f_T(A)$ less than or equal to x .

Let $h(x+1) := \langle 1, 0 \rangle$.

That is, Cases 1 and 2 correspond to the situations $\text{Pr}_T(\ulcorner f_T(A) \urcorner) \preceq \sigma$ and $\sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner)$, respectively.

After that, we define the value of $h(x+1)$ as follows:

$$h(x+1) := \begin{cases} a & \text{if } h(x) \sqsubset' a \text{ \& } x \text{ is a } T\text{-proof of } \neg\lambda(\bar{a}); \\ h(x) & \text{otherwise.} \end{cases}$$

The definition of h is hereby completed.

We introduce three lemmas. The first lemma concerns the general properties of the function $h(x)$, the formula $\lambda(x)$, and the arithmetical interpretation f_T .

LEMMA 3.15. *Let $a, b \in W'$.*

1. *If $a \neq b$, then $PA \vdash \lambda(\bar{a}) \rightarrow \neg\lambda(\bar{b})$,*
2. *$PA \vdash h(x) = \bar{a} \rightarrow \lambda(\bar{a}) \vee \bigvee_{a \sqsubset' c} \lambda(\bar{c})$,*
3. *$PA \vdash \text{Pr}_T(\ulcorner f_T(A) \urcorner) \preceq \sigma \leftrightarrow \bigvee_{\pi_0(c)=0} \lambda(\bar{c})$,*
4. *$PA \vdash \sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner) \leftrightarrow \bigvee_{\pi_0(c)=1} \lambda(\bar{c})$,*
5. *If $a \sqsubset' b$, then $PA \vdash \lambda(\bar{a}) \rightarrow \neg\text{Pr}_T(\ulcorner \neg\lambda(\bar{b}) \urcorner)$,*
6. *If $\pi_1(a) \geq 1$, then $PA \vdash \lambda(\bar{a}) \rightarrow \text{Pr}_T(\ulcorner \neg\lambda(\bar{a}) \urcorner)$,*
7. *If $\pi_1(a) \geq 1$, then $PA \vdash \lambda(\bar{a}) \rightarrow \text{Pr}_T(\ulcorner \bigvee_{a \sqsubset' c} \lambda(\bar{c}) \urcorner)$.*

PROOF. Except the implications \leftarrow in Clauses 3 and 4, these statements are proved in a similar way as in the usual proof of Solovay's arithmetical completeness theorem (cf. [1, 2, 6, 16]).

For the implication \leftarrow in Clause 3, we argue in PA : Suppose that $\lambda(\bar{c})$ holds and $\pi_0(c) = 0$. Then, $h(k) = c$ for some k . If $\neg\sigma$ and $\neg\text{Pr}_T(\ulcorner f_T(A) \urcorner)$ hold, then $\forall x h(x) = 0$ holds, a contradiction. Hence, $\text{Pr}_T(\ulcorner f_T(A) \urcorner) \preceq \sigma$ or $\sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner)$ holds. By Clause 1, we have $\neg\bigvee_{\pi_0(d)=1} \lambda(\bar{d})$. Then, by the implication \rightarrow in Clause 4, $\sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner)$ does not hold. Therefore, $\text{Pr}_T(\ulcorner f_T(A) \urcorner) \preceq \sigma$ holds.

The implication \leftarrow in Clause 4 is also proved similarly. ■

The second lemma states that the satisfaction relation for $a \in W'$ with $\pi_1(a) \geq 1$ is embedded into PA .

LEMMA 3.16. *Let $a \in W'$ with $\pi_1(a) \geq 1$ and let B be a modal formula.*

1. *If $a \Vdash' B$, then $PA \vdash \lambda(\bar{a}) \rightarrow f_T(B)$.*
2. *If $a \nVdash' B$, then $PA \vdash \lambda(\bar{a}) \rightarrow \neg f_T(B)$.*

PROOF. Clauses 1 and 2 are proved by induction on the construction of B simultaneously as in the usual proof of Solovay's theorem. ■

The third lemma concerns the satisfaction relation for the element $\langle 1, 1 \rangle$.

LEMMA 3.17. *Let B be any subformula of $\Box A$.*

1. *If $\langle 1, 1 \rangle \Vdash' B$, then $PA \vdash \lambda(\overline{\langle 1, 0 \rangle}) \rightarrow f_T(B)$,*
2. *If B is $\Box C$ and $\langle 1, 1 \rangle \Vdash' \Box C$, then $PA \vdash \sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner) \rightarrow f_T(\Box C)$,*
3. *If $\langle 1, 1 \rangle \nVdash' B$, then $PA \vdash \lambda(\overline{\langle 1, 0 \rangle}) \rightarrow \neg f_T(B)$.*

PROOF. We prove Clauses 1, 2 and 3 by induction on the construction of $B \in \text{Sub}(\Box A)$ simultaneously. We only give a proof of the case that B is $\Box C$.

1 and 2: Suppose $\langle 1, 1 \rangle \Vdash' \Box C$. Since $\Box C \in \text{Sub}(\Box A)$, we have $\langle 1, 1 \rangle \Vdash' \Box C \rightarrow C$ because $\langle 1, 1 \rangle \Vdash' \bigwedge \text{Rfn}(\Box A)$. Then, $\langle 1, 1 \rangle \Vdash' C$. Hence, for all $a \in W'$ with $\pi_0(a) = 1$ and $\pi_1(a) \geq 1$, we have $a \Vdash' C$. By Lemma 3.16.1, $PA \vdash \bigvee_{\substack{\pi_0(a)=1 \\ \pi_1(a) \geq 1}} \lambda(\bar{a}) \rightarrow f_T(C)$. Also, by the induction hypothesis, $PA \vdash \lambda(\overline{\langle 1, 0 \rangle}) \rightarrow f_T(C)$. Thus,

$$PA \vdash \bigvee_{\pi_0(a)=1} \lambda(\bar{a}) \rightarrow f_T(C).$$

By Lemma 3.15.4,

$$PA \vdash \sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner) \rightarrow f_T(C).$$

Then,

$$PA \vdash \text{Pr}_T(\ulcorner \sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner) \urcorner) \rightarrow f_T(\Box C).$$

Since $\sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner)$ is a Σ_1 sentence, we have

$$PA \vdash \sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner) \rightarrow f_T(\Box C).$$

This completes the proof of Clause 2. $PA \vdash \lambda(\overline{\langle 1, 0 \rangle}) \rightarrow \sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner)$ by Lemma 3.15.4, and hence Clause 1 also holds.

3: Suppose $\langle 1, 1 \rangle \nVdash' \Box C$. Then, $a \nVdash' C$ for some $a \in W'$ with $\langle 1, 1 \rangle \sqsubset' a$. By Lemma 3.16.2, we have

$$PA \vdash \lambda(\bar{a}) \rightarrow \neg f_T(C).$$

Then,

$$PA \vdash \neg \text{Pr}_T(\ulcorner \neg \lambda(\bar{a}) \urcorner) \rightarrow \neg f_T(\Box C).$$

Since $\langle 1, 0 \rangle \sqsubset' a$, by Lemma 3.15.5,

$$\text{PA} \vdash \lambda(\overline{\langle 1, 0 \rangle}) \rightarrow \neg \text{Pr}_T(\ulcorner \neg \lambda(\bar{a}) \urcorner).$$

Therefore, $\text{PA} \vdash \lambda(\overline{\langle 1, 0 \rangle}) \rightarrow \neg f_T(\Box C)$. ■

We are ready to show the required two statements: (a) $\text{PA} \vdash \sigma \rightarrow f_T(\Box A)$ and (b) $\text{PA} + \text{Con}_T^{c(A)+1} \vdash f_T(\Box A) \rightarrow \sigma$.

(a): Since $\langle 1, 1 \rangle \Vdash' \Box A$, by Lemma 3.17.2,

$$\text{PA} \vdash \sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner) \rightarrow f_T(\Box A).$$

Since

$$\text{PA} \vdash \sigma \wedge \neg f_T(\Box A) \rightarrow \sigma \prec \text{Pr}_T(\ulcorner f_T(A) \urcorner),$$

we have

$$\text{PA} \vdash \sigma \wedge \neg f_T(\Box A) \rightarrow f_T(\Box A),$$

and hence we obtain $\text{PA} \vdash \sigma \rightarrow f_T(\Box A)$.

(b): Since $\langle 0, 1 \rangle \Vdash' \Box^{c(A)+1} \perp$, for all $a \in W'$ with $\pi_0(a) = 0$ and $\pi_1(a) \geq 1$, we have $a \Vdash' \Box^{c(A)+1} \perp$. By Lemma 3.16.1,

$$\text{PA} \vdash \bigvee_{\substack{\pi_0(a)=0 \\ \pi_1(a) \geq 1}} \lambda(\bar{a}) \rightarrow \neg \text{Con}_T^{c(A)+1},$$

and thus

$$\text{PA} + \text{Con}_T^{c(A)+1} \vdash \bigwedge_{\substack{\pi_0(a)=0 \\ \pi_1(a) \geq 1}} \neg \lambda(\bar{a}). \quad (\text{IV})$$

Since

$$\text{PA} \vdash f_T(\Box A) \wedge \neg \sigma \rightarrow \text{Pr}_T(\ulcorner f_T(A) \urcorner) \preceq \sigma,$$

by Lemma 3.15.3, we have

$$\text{PA} \vdash f_T(\Box A) \wedge \neg \sigma \rightarrow \bigvee_{\pi_0(a)=0} \lambda(\bar{a}).$$

By combining this with (IV), we obtain

$$\text{PA} + \text{Con}_T^{c(A)+1} \vdash f_T(\Box A) \wedge \neg \sigma \rightarrow \lambda(\overline{\langle 0, 0 \rangle}). \quad (\text{V})$$

Since $\langle 0, 1 \rangle \Vdash' A$, by Lemma 3.16.2, $\text{PA} \vdash \lambda(\overline{\langle 0, 1 \rangle}) \rightarrow \neg f_T(A)$. Then,

$$\text{PA} \vdash \neg \text{Pr}_T(\ulcorner \neg \lambda(\overline{\langle 0, 1 \rangle}) \urcorner) \rightarrow \neg f_T(\Box A).$$

Since $\langle 0, 0 \rangle \sqsubset' \langle 0, 1 \rangle$, by Lemma 3.15.5,

$$\text{PA} \vdash \lambda(\overline{\langle 0, 0 \rangle}) \rightarrow \neg \text{Pr}_T(\ulcorner \neg \lambda(\overline{\langle 0, 1 \rangle}) \urcorner),$$

and hence we have $\text{PA} \vdash \lambda(\overline{\langle 0, 0 \rangle}) \rightarrow \neg f_T(\Box A)$. From this with (V), we obtain $\text{PA} + \text{Con}_T^{\mathfrak{c}(A)+1} \vdash f_T(\Box A) \wedge \neg \sigma \rightarrow \neg f_T(\Box A)$. We conclude $\text{PA} + \text{Con}_T^{\mathfrak{c}(A)+1} \vdash f_T(\Box A) \rightarrow \sigma$. ■

If the height of T is larger than $\mathfrak{c}(A)$, then the converse of Theorem 3.14 also holds.

PROPOSITION 3.18. *Let A be any modal formula. Suppose that the height of T is larger than $\mathfrak{c}(A)$ and, for any Σ_1 sentence σ , there exists an arithmetical interpretation f such that $\text{PA} \vdash \sigma \rightarrow f_T(\Box A)$ and $\text{PA} + \text{Con}_T^{\mathfrak{c}(A)+1} \vdash f_T(\Box A) \rightarrow \sigma$. Then, A is nontrifling.*

PROOF. Since $0 = 1$ is a Σ_1 sentence, there exists an arithmetical interpretation f such that $\text{PA} + \text{Con}_T^{\mathfrak{c}(A)+1} \vdash f_T(\Box A) \rightarrow 0 = 1$. Since $\mathbb{N} \models \text{Con}_T^{\mathfrak{c}(A)+1}$, we have $\mathbb{N} \models \neg f_T(\Box A)$. Thus, $T \not\models f_T(A)$. By Fact 3.3.1, $\mathbf{GL} \not\models A$.

Also, since $0 = 0$ is a Σ_1 sentence, there exists an arithmetical interpretation g such that $\text{PA} \vdash 0 = 0 \rightarrow g_T(\Box A)$. Then, $\text{PA} \vdash g_T(\Box A)$. Since $\mathbb{N} \models \text{Con}_T^{\mathfrak{c}(A)+1}$, $\text{PA} \not\models \neg \text{Con}_T^{\mathfrak{c}(A)+1}$, and thus $\text{PA} \not\models \text{Con}_T^{\mathfrak{c}(A)+1} \rightarrow \neg g_T(\Box A)$. By Fact 3.3.1, $\mathbf{GL} \not\models \Diamond^{\mathfrak{c}(A)+1} \top \rightarrow \neg \Box A$. By Lemma 3.6, $\mathbf{GL}_\omega \not\models \neg \Box A$. Then, by Proposition 3.11, A is nontrifling. ■

For any classical propositional formula A , we have $\mathfrak{c}(A) = 0$. So Proposition 3.13 shows that Theorem 3.14 is actually an extension of Theorem 2.3.

For a variable v of first-order logic, an v -arithmetical interpretation f is a mapping where for each propositional variable p , $f(p)$ is an \mathcal{L}_A -formula $\varphi(v)$ with only the free variable v . Each v -arithmetical interpretation f is uniquely extended to the mapping f_T from the set of all modal formulas to a set of \mathcal{L}_A -formulas with at most the free variable v as the usual arithmetical interpretations with the clause $f_T(\Box A)(v) \equiv \text{Pr}_T(\ulcorner f_T(A)(\dot{v}) \urcorner)$. By tracing the proof of Theorem 3.14 entirely using the function $h(x, v)$ and the formula $\lambda(x, v)$, the following parameterized version of Theorem 3.14 is also proved.

THEOREM 3.19. *For any modal formula A , if A is nontrifling, then for any Σ_1 formula $\sigma(v)$, there exists an v -arithmetical interpretation f such that*

1. $\text{PA} \vdash \forall v (\sigma(v) \rightarrow f_T(\Box A)(v))$, and
2. $\text{PA} + \text{Con}_T^{\mathfrak{c}(A)+1} \vdash \forall v (f_T(\Box A)(v) \rightarrow \sigma(v))$.

From Theorem 3.19, we obtain an extension of Theorem 2.5 to the framework of modal logic.

THEOREM 3.20. *Let A be any modal formula A . If A is nontrifling and the height of T is larger than $\mathfrak{c}(A)$, then for any c.e. set X , there exists an v -arithmetical interpretation f such that $f_T(A)(v)$ weakly represents X in T .*

PROOF. Let $\sigma(v)$ be a Σ_1 formula defining X over \mathbb{N} . By Theorem 3.19, there exists an v -arithmetical interpretation f such that

1. $\text{PA} \vdash \forall v (\sigma(v) \rightarrow f_T(\Box A)(v))$, and
2. $\text{PA} + \text{Con}_T^{\mathfrak{c}(A)+1} \vdash \forall v (f_T(\Box A)(v) \rightarrow \sigma(v))$.

Since $\mathbb{N} \models \text{Con}_T^{\mathfrak{c}(A)+1}$, we have $\mathbb{N} \models \forall v (\sigma(v) \leftrightarrow f_T(\Box A)(v))$. Then, for any $k \in \omega$, $k \in X$ if and only if $T \vdash f_T(A)(\bar{k})$. This means that $f_T(A)(v)$ weakly represents X in T . ■

3.3. A modal version of the FGH Theorem for Theories with Finite Heights

Notice that Theorem 3.20 requires the assumption that the height of T is larger than $\mathfrak{c}(A)$. In this subsection, we investigate variations of Theorems 3.14 and 3.20, keeping in mind theories with finite heights.

It is already proved in Proposition 3.11 that for any natural number s with $s > \mathfrak{c}(A)$, A is nontrifling if and only if $\mathbf{GL} + \{\neg F_s\} \not\vdash \Box A$ and $\mathbf{GL} + \{\neg F_s\} \not\vdash \neg \Box A$. On the other hand, even for a natural number s with $s \leq \mathfrak{c}(A)$, we prove the following version of the FGH theorem concerning the condition “ $\mathbf{GL} + \{\neg F_s\} \not\vdash \Box A$ and $\mathbf{GL} + \{\neg F_s\} \not\vdash \neg \Box A$ ”.

THEOREM 3.21. *For any modal formula A , if $\mathbf{GL} + \{\neg F_s\} \not\vdash \Box A$ and $\mathbf{GL} + \{\neg F_s\} \not\vdash \neg \Box A$, then for any Σ_1 sentence σ , there exists an arithmetical interpretation f such that*

$$\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \sigma \leftrightarrow f_T(\Box A).$$

PROOF. Suppose $\mathbf{GL} + \{\neg F_s\} \not\vdash \Box A$ and $\mathbf{GL} + \{\neg F_s\} \not\vdash \neg \Box A$, that is, $\mathbf{GL} \not\vdash \Box^{s+1} \perp \wedge \Diamond^s \top \rightarrow \Box A$ and $\mathbf{GL} \not\vdash \Box^{s+1} \perp \wedge \Diamond^s \top \rightarrow \neg \Box A$. By the Kripke completeness of \mathbf{GL} , there exist \mathbf{GL} -models $M_0 = \langle W_0, \sqsubset_0, r_0, \Vdash_0 \rangle$ and $M_1 = \langle W_1, \sqsubset_1, r_1, \Vdash_1 \rangle$ such that $r_0 \Vdash_0 \Box^{s+1} \perp \wedge \Diamond^s \top \wedge \neg \Box A$ and $r_1 \Vdash_1 \Box^{s+1} \perp \wedge \Diamond^s \top \wedge \Box A$. Let n_0 and n_1 be the cardinalities of W_0 and W_1 respectively, and we may assume that

- $W_0 = \{\langle 0, 1 \rangle, \langle 0, 2 \rangle, \dots, \langle 0, n_0 \rangle\}$, $r_0 = \langle 0, 1 \rangle$, and
- $W_1 = \{\langle 1, 1 \rangle, \langle 1, 2 \rangle, \dots, \langle 1, n_1 \rangle\}$, $r_1 = \langle 1, 1 \rangle$.

Let $M' = \langle W', \sqsubset', 0, \Vdash' \rangle$ be the **GL**-model obtained from M_0 and M_1 as in the proof of Theorem 3.14. Then, $\langle 0, 1 \rangle \Vdash' \Box^{s+1} \perp \wedge \Diamond^s \top \wedge \neg \Box A$ and $\langle 1, 1 \rangle \Vdash' \Box^{s+1} \perp \wedge \Diamond^s \top \wedge \Box A$.

We can define a function h , a formula $\lambda(x)$, and an arithmetical interpretation f from the **GL**-model M' in the same way as in the proof of Theorem 3.14. Then, the same results as Lemmas 3.15, 3.16 and 3.17 can also be proved.

For each $a \in W'$ with $\pi_1(a) > 1$, since $a \Vdash' \Box^s \perp$, we have that PA proves $\lambda(\bar{a}) \rightarrow \neg \text{Con}_T^s$, and thus

$$\text{PA} + \text{Con}_T^s \vdash \bigwedge_{\pi_1(a) > 1} \neg \lambda(\bar{a}).$$

Also, since $\langle i, 1 \rangle \Vdash' \Diamond^s \top$, we have $\text{PA} \vdash \lambda(\overline{\langle i, 1 \rangle}) \rightarrow \text{Con}_T^s$. Then, PA proves $\neg \text{Pr}_T(\neg \lambda(\overline{\langle i, 1 \rangle})) \rightarrow \text{Con}_T^{s+1}$. Also, $\text{PA} \vdash \lambda(\overline{\langle i, 0 \rangle}) \rightarrow \neg \text{Pr}_T(\neg \lambda(\overline{\langle i, 1 \rangle}))$ because $\langle i, 0 \rangle \sqsubset' \langle i, 1 \rangle$, and hence $\text{PA} \vdash \lambda(\overline{\langle i, 0 \rangle}) \rightarrow \text{Con}_T^{s+1}$. Therefore, we obtain

$$\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \bigwedge_{\pi_1(a) \neq 1} \neg \lambda(\bar{a}). \quad (\text{VI})$$

We prove $\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \sigma \leftrightarrow f_T(\Box A)$.

(\rightarrow): Since $\text{PA} \vdash \sigma \prec \text{Pr}_T(\neg f_T(A)) \rightarrow \bigvee_{\pi_0(a)=1} \lambda(\bar{a})$, from (VI), we have

$$\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \sigma \prec \text{Pr}_T(\neg f_T(A)) \rightarrow \lambda(\overline{\langle 1, 1 \rangle}).$$

Since $\langle 1, 1 \rangle \Vdash' \Box A$, we have $\text{PA} \vdash \lambda(\overline{\langle 1, 1 \rangle}) \rightarrow f_T(\Box A)$. Hence,

$$\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \sigma \prec \text{Pr}_T(\neg f_T(A)) \rightarrow f_T(\Box A).$$

Here $\text{PA} \vdash \sigma \wedge \neg f_T(\Box A) \rightarrow \sigma \prec \text{Pr}_T(\neg f_T(A))$, and thus

$$\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \sigma \wedge \neg f_T(\Box A) \rightarrow f_T(\Box A).$$

We obtain $\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \sigma \rightarrow f_T(\Box A)$.

(\leftarrow): Since $\text{PA} \vdash \text{Pr}_T(\neg f_T(A)) \preceq \sigma \rightarrow \bigvee_{\pi_0(a)=0} \lambda(\bar{a})$, from (VI),

$$\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \text{Pr}_T(\neg f_T(A)) \preceq \sigma \rightarrow \lambda(\overline{\langle 0, 1 \rangle}).$$

Since $\langle 0, 1 \rangle \Vdash' \Box A$, $\text{PA} \vdash \lambda(\overline{\langle 0, 1 \rangle}) \rightarrow \neg f_T(\Box A)$. Therefore,

$$\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash \text{Pr}_T(\neg f_T(A)) \preceq \sigma \rightarrow \neg f_T(\Box A),$$

and hence

$$\text{PA} + \neg \text{Con}_T^{s+1} + \text{Con}_T^s \vdash f_T(\Box A) \wedge \neg \sigma \rightarrow \neg f_T(\Box A).$$

We conclude $\text{PA} + \neg\text{Con}_T^{s+1} + \text{Con}_T^s \vdash f_T(\Box A) \rightarrow \sigma$. \blacksquare

If the height of T is larger than or equal to s , then the converse of Theorem 3.21 also holds.

PROPOSITION 3.22. *Suppose that the height of T is larger than or equal to s . For any modal formula A , if for any Σ_1 sentence σ , there exists an arithmetical interpretation f such that $\text{PA} + \neg\text{Con}_T^{s+1} + \text{Con}_T^s \vdash \sigma \leftrightarrow f_T(\Box A)$, then $\mathbf{GL} + \{\neg F_s\} \not\vdash \Box A$ and $\mathbf{GL} + \{\neg F_s\} \not\vdash \neg\Box A$.*

PROOF. If the height of T is larger than s , then $T \not\vdash \neg\text{Con}_T^s$. By Löb's theorem, $T \not\vdash \neg\text{Con}_T^{s+1} \rightarrow \neg\text{Con}_T^s$. Thus, the theory $T + \neg\text{Con}_T^{s+1} + \text{Con}_T^s$ is consistent. Also, if the height of T is exactly s , then $\mathbb{N} \models \neg\text{Con}_T^{s+1} \wedge \text{Con}_T^s$. Thus, $\text{PA} + \neg\text{Con}_T^{s+1} + \text{Con}_T^s$ is consistent. In either case, the theory $\text{PA} + \neg\text{Con}_T^{s+1} + \text{Con}_T^s$ is consistent.

Let σ be any Σ_1 sentence independent of $\text{PA} + \neg\text{Con}_T^{s+1} + \text{Con}_T^s$. Then, we have an arithmetical interpretation f such that

$$\text{PA} + \neg\text{Con}_T^{s+1} + \text{Con}_T^s \vdash \sigma \leftrightarrow f_T(\Box A).$$

Then, $\text{PA} + \neg\text{Con}_T^{s+1} + \text{Con}_T^s$ proves neither $f_T(\Box A)$ nor $\neg f_T(\Box A)$. By Fact 3.3.1, \mathbf{GL} proves neither $\neg F_s \rightarrow \Box A$ nor $\neg F_s \rightarrow \neg\Box A$. \blacksquare

In the same way as in the proof of Theorem 3.21, the parameterized version of Theorem 3.21 can also be proved, and hence we obtain the following theorem.

THEOREM 3.23. *Suppose that the height of T is s . For any modal formula A , if $\mathbf{GL} + \{\neg F_s\} \not\vdash \Box A$ and $\mathbf{GL} + \{\neg F_s\} \not\vdash \neg\Box A$, then for any c.e. set X , there exists a v -arithmetical interpretation f such that $f_T(A)(v)$ weakly represents X in T .*

PROOF. Let $\sigma(v)$ be a Σ_1 formula defining X over \mathbb{N} . Then, from the parameterized version of Theorem 3.21, there exists a v -arithmetical interpretation f such that

$$\text{PA} + \neg\text{Con}_T^{s+1} + \text{Con}_T^s \vdash \forall v (\sigma(v) \leftrightarrow f_T(\Box A)(v)).$$

Since the height of T is s , $\mathbb{N} \models \neg\text{Con}_T^{s+1} \wedge \text{Con}_T^s$. Then, we have that $\mathbb{N} \models \forall v (\sigma(v) \leftrightarrow f_T(\Box A)(v))$. It follows that $f_T(A)(v)$ weakly represents X in T . \blacksquare

We close this section with two problems. In Section 2, we proved Theorem 2.3 by using Lemma 2.4. On the other hand, in this section, we proved Theorem 3.14 that is a modal extension of Theorem 2.3 without using modal

version of Lemma 2.4. If a modal version of Lemma 2.4 is proved, then a proof of Theorem 3.14, as well as the proof of our Theorem 2.3 in Section 2, would be substantially simplified.

PROBLEM 3.24. Can we prove a modal version of Lemma 2.4?

In Section 2, we proved Theorem 2.6 that is an improvement of Theorem 2.3. Then, it is natural to ask the following problem.

PROBLEM 3.25. Can we extend Theorem 2.6 to the framework of modal propositional logic?

4. Rosser-Type FGH Theorems

In this section, we investigate some variations of Rosser-type FGH theorems. Recall that $\text{Proof}_T(x, y)$ is a natural formula saying that y is a T -proof of x and $\text{Pr}_T(x)$ is of the form $\exists y \text{Proof}_T(x, y)$. Besides $\text{Proof}_T(x, y)$, we consider formulas that witness $\text{Pr}_T(x)$. We say a formula $\text{Prf}_T(x, y)$ is a *proof predicate* of T if $\text{Prf}_T(x, y)$ satisfies the following conditions:

1. $\text{Prf}_T(x, y)$ is a primitive recursive formula,
2. $\text{PA} \vdash \forall x (\text{Pr}_T(x) \leftrightarrow \exists y \text{Prf}_T(x, y))$,
3. $\text{PA} \vdash \forall x \forall x' \forall y (\text{Prf}_T(x, y) \wedge \text{Prf}_T(x', y) \rightarrow x = x')$.

For each proof predicate $\text{Prf}_T(x, y)$ of T , the Σ_1 formula $\text{Pr}_T^R(x)$ defined by

$$\text{Pr}_T^R(x) := \exists y (\text{Prf}_T(x, y) \wedge \forall z < y \neg \text{Prf}_T(\dot{\neg}x, z))$$

is called the *Rosser provability predicate* of $\text{Prf}_T(x, y)$ or a Rosser provability predicate of T . Here $\dot{\neg}x$ is a primitive recursive term corresponding to a primitive recursive function calculating the Gödel number of $\neg\varphi$ from the Gödel number of an \mathcal{L}_A -formula φ such that PA proves $\dot{\neg}\ulcorner\psi\urcorner = \ulcorner\neg\psi\urcorner$ for each \mathcal{L}_A -formula ψ . In view of witness comparison, we also introduce an auxiliary Σ_1 formula $\text{Pr}_T^A(\dot{\neg}x)$ as follows:

$$\text{Pr}_T^A(\dot{\neg}x) := \exists y (\text{Prf}_T(\dot{\neg}x, y) \wedge \forall z \leq y \neg \text{Prf}_T(x, z)).$$

Then, for any \mathcal{L}_A -formula φ , PA proves $\neg(\text{Pr}_T^R(\ulcorner\varphi\urcorner) \wedge \text{Pr}_T^A(\ulcorner\neg\varphi\urcorner))$ and $\text{Pr}_T(\ulcorner\varphi\urcorner) \vee \text{Pr}_T(\ulcorner\neg\varphi\urcorner) \rightarrow \text{Pr}_T^R(\ulcorner\varphi\urcorner) \vee \text{Pr}_T^A(\ulcorner\neg\varphi\urcorner)$.

In a study of Rosser-type Henkin sentences, the following result was proved.

FACT 4.1. (Kurahashi [9, Theorem 3.5]) For any Σ_1 sentence σ , the following are equivalent:

1. There exists a Σ_1 sentence σ' such that $PA \vdash \neg(\sigma \wedge \sigma')$ and $PA \vdash \Pr_T(\ulcorner \sigma \urcorner) \vee \Pr_T(\ulcorner \sigma' \urcorner) \rightarrow \sigma \vee \sigma'$.
2. There exists a Rosser provability predicate $\Pr_T^R(x)$ of T such that $PA \vdash \sigma \leftrightarrow \Pr_T^R(\ulcorner \sigma \urcorner)$.

This fact can be seen as a kind of the FGH theorem for Rosser provability predicates because it deals with the equivalence $\sigma \leftrightarrow \Pr_T^R(\ulcorner \varphi \urcorner)$ where φ is in particular σ . Inspired by this fact, we prove the following theorem.

THEOREM 4.2. For any Σ_1 sentences σ_0 and σ_1 , the following are equivalent:

1. $PA \vdash \neg(\sigma_0 \wedge \sigma_1)$ and $PA \vdash \neg \text{Con}_T \rightarrow \sigma_0 \vee \sigma_1$.
2. There are a Rosser provability predicate $\Pr_T^R(x)$ of T and an \mathcal{L}_A -sentence φ such that

$$PA \vdash \sigma_0 \leftrightarrow \Pr_T^R(\ulcorner \varphi \urcorner) \text{ and } PA \vdash \sigma_1 \leftrightarrow \Pr_T^R(\ulcorner \neg \varphi \urcorner).$$

PROOF. (2 \Rightarrow 1): This implication follows from the properties of witness comparison formulas.

(1 \Rightarrow 2): We may assume that σ_0 and σ_1 are of the forms $\exists x \tau_0(x)$ and $\exists x \tau_1(x)$ for some Δ_0 formulas $\tau_0(x)$ and $\tau_1(x)$, respectively. By the Fixed Point Lemma, let φ be a Σ_1 sentence satisfying the following equivalence:

$$PA \vdash \varphi \leftrightarrow \exists y [(\text{Proof}_T(\ulcorner \neg \varphi \urcorner, y) \vee \tau_0(y)) \wedge \forall z \leq y (\neg \text{Proof}_T(\ulcorner \neg \varphi \urcorner, z) \wedge \neg \tau_1(z))].$$

Also, let φ^* be the Σ_1 sentence

$$\exists z [(\text{Proof}_T(\ulcorner \varphi \urcorner, z) \vee \tau_1(z)) \wedge \forall y < z (\neg \text{Proof}_T(\ulcorner \neg \varphi \urcorner, y) \wedge \neg \tau_0(y))].$$

LEMMA 4.3. $PA \vdash \Pr_T(\ulcorner \varphi \urcorner) \vee \Pr_T(\ulcorner \neg \varphi \urcorner) \rightarrow \sigma_0 \vee \sigma_1$.

PROOF. Let $\text{Prov}_T^R(x)$ be the Rosser provability predicate of $\text{Proof}_T(x, y)$. Then, $PA \vdash \Pr_T(\ulcorner \varphi \urcorner) \vee \Pr_T(\ulcorner \neg \varphi \urcorner) \rightarrow \text{Prov}_T^R(\ulcorner \varphi \urcorner) \vee \text{Prov}_T^R(\ulcorner \neg \varphi \urcorner)$. By the definition of φ^* ,

$$PA \vdash \text{Prov}_T^R(\ulcorner \varphi \urcorner) \wedge \neg \sigma_0 \rightarrow \varphi^*.$$

Then,

$$PA \vdash \text{Prov}_T^R(\ulcorner \varphi \urcorner) \wedge \neg \sigma_0 \rightarrow \Pr_T(\ulcorner \varphi^* \urcorner)$$

because φ^* is a Σ_1 sentence. Since $PA \vdash \neg(\varphi \wedge \varphi^*)$, we have

$$PA \vdash \text{Prov}_T^R(\ulcorner \varphi \urcorner) \wedge \neg \sigma_0 \rightarrow \neg \text{Con}_T. \quad (\text{VII})$$

By the choice of φ ,

$$\text{PA} \vdash \text{Prov}_T^R(\ulcorner \neg \varphi \urcorner) \wedge \neg \sigma_1 \rightarrow \varphi,$$

and we also obtain

$$\text{PA} \vdash \text{Prov}_T^R(\ulcorner \neg \varphi \urcorner) \wedge \neg \sigma_1 \rightarrow \neg \text{Con}_T.$$

From this with (VII),

$$\text{PA} + \text{Con}_T \vdash \text{Prov}_T^R(\ulcorner \varphi \urcorner) \vee \text{Prov}_T^R(\ulcorner \neg \varphi \urcorner) \rightarrow \sigma_0 \vee \sigma_1.$$

Thus,

$$\text{PA} + \text{Con}_T \vdash \text{Pr}_T(\ulcorner \varphi \urcorner) \vee \text{Pr}_T(\ulcorner \neg \varphi \urcorner) \rightarrow \sigma_0 \vee \sigma_1.$$

Since $\text{PA} + \neg \text{Con}_T \vdash \sigma_0 \vee \sigma_1$, we conclude

$$\text{PA} \vdash \text{Pr}_T(\ulcorner \varphi \urcorner) \vee \text{Pr}_T(\ulcorner \neg \varphi \urcorner) \rightarrow \sigma_0 \vee \sigma_1.$$

■

We proceed with the main proof. We recursively define a primitive recursive function h and an increasing sequence $\langle k_i \rangle_{i \in \omega}$ of natural numbers simultaneously by referring to T -proofs in stages. The function h will be defined to output all theorems of T and the Rosser predicate of the proof predicate $x = h(y)$ will be a required one. At the beginning of Stage m , the values of k_0, \dots, k_{m-1}, k_m and $h(0), \dots, h(k_{m-1})$ have already been defined.

Here we give our definition of the function h . In the definition, we identify each \mathcal{L}_A -formula with its Gödel number. First, let $k_0 = 0$.

At Stage m ,

- If m is not a T -proof of any \mathcal{L}_A -formula, then let $k_{m+1} := k_m$ and go to Stage $m + 1$.
- If m is a T -proof of an \mathcal{L}_A -formula ξ which is neither φ nor $\neg \varphi$, then let $k_{m+1} := k_m + 1$ and $h(k_m) := \xi$, and go to the next stage.
- If m is a T -proof of an \mathcal{L}_A -sentence ξ which is either φ or $\neg \varphi$, and h already outputs at least one of φ or $\neg \varphi$ before Stage m , then let $k_{m+1} := k_m + 1$ and $h(k_m) := \xi$, and go to Stage $m + 1$.
- If m is a T -proof of either φ or $\neg \varphi$, and f does not output φ and $\neg \varphi$ before Stage m , then we distinguish the following three cases:
 - (a) If $\exists z \leq m \tau_0(z)$ holds, then let $k_{m+1} := k_m + 1$ and $h(k_m) := \varphi$.
 - (b) If $\exists z \leq m \tau_1(z)$ holds, then let $k_{m+1} := k_m + 1$ and $h(k_m) := \neg \varphi$.
 - (c) Otherwise, let $k_{m+1} := k_m$.

Go to the next Stage $m + 1$.

This completes our definition the function h . Since $\text{PA} \vdash \neg(\sigma_0 \wedge \sigma_1)$, we have that PA proves that there is no m satisfying both $\exists z \leq m \tau_0(z)$ and $\exists z \leq m \tau_1(z)$. Thus Cases (a) and (b) in the definition of h are mutually exclusive.

We show that the formula $x = h(y)$ is a proof predicate of T . It suffices to show the following lemma:

LEMMA 4.4. $\text{PA} \vdash \forall x (\text{Pr}_T(x) \leftrightarrow \exists y (x = h(y)))$.

PROOF. We argue in PA .

(\rightarrow): Suppose that ξ is provable in T .

If ξ is neither φ nor $\neg\varphi$, then for a T -proof m of ξ , $h(k_m) = \xi$.

If ξ is either φ or $\neg\varphi$, then by Lemma 4.3, $\sigma_0 \vee \sigma_1$ holds. However, σ_0 and σ_1 cannot be true at the same time. We show that ξ is output by h . We distinguish the following two cases:

- Case 1: σ_0 holds.

Let n be the least number such that $\tau_0(n)$ holds, and let m be the least number such that $m \geq n$ and m is a T -proof of either φ or $\neg\varphi$. Then, $h(k_m) = \varphi$ by the definition of h . If ξ is φ , we are done. If ξ is $\neg\varphi$, then let $m' > m$ be a T -proof of $\neg\varphi$. Such an m' exists because $\neg\varphi$ has infinitely many T -proofs. Since h already outputs φ before Stage m' , by the definition of h , $h(k_{m'}) = \neg\varphi$.

- Case 2: σ_1 holds.

It is proved that h outputs ξ as in the proof of Case 1 by considering the least number n such that $\tau_1(n)$ holds.

(\leftarrow): Suppose that ξ is output by h . If ξ is neither φ nor $\neg\varphi$, then $h(k_m) = \xi$ implies that m is a T -proof of ξ , and so ξ is provable in T .

If ξ is φ or $\neg\varphi$, then let m be the least number such that $h(k_m)$ is either φ or $\neg\varphi$. We show that ξ is provable in T . We distinguish the following four cases:

- Case 1: ξ is φ and $h(k_m) = \varphi$.

By the definition of h , $\exists z \leq m \tau_0(z)$ holds. Then, σ_0 and $\neg\sigma_1$ hold. Suppose that φ is not provable in T , then φ holds by the definition. Since φ is a Σ_1 sentence, it is provable in T , a contradiction. Therefore, φ is provable in T .

- Case 2: ξ is φ and $h(k_m) = \neg\varphi$.

Then, $h(k_{m'}) = \varphi$ for some $m' > m$. Since $\neg\varphi$ is already output before Stage m' , we find that m' is a T -proof of φ , and hence φ is provable in T .

- Case 3: ξ is $\neg\varphi$ and $h(k_m) = \varphi$.

Then, for some T -proof $m' > m$ of $\neg\varphi$, $h(k_{m'}) = \neg\varphi$. Thus, $\neg\varphi$ is provable in T .

- Case 4: ξ is $\neg\varphi$ and $h(k_m) = \neg\varphi$.

Then, $\exists z \leq m \tau_1(z)$ holds, and so σ_1 and $\neg\sigma_0$ hold. Suppose that $\neg\varphi$ is not provable in T . Then, φ^* holds, and is provable in T . Since $\varphi^* \rightarrow \neg\varphi$ is T -provable, $\neg\varphi$ is also T -provable. This is a contradiction. Therefore, $\neg\varphi$ is provable in T . ■

We resume the main proof. Let $\text{Pr}_h^R(x)$ be the Rosser provability predicate of the proof predicate $x = h(y)$ of T . Finally, we show that PA proves $\sigma_0 \leftrightarrow \text{Pr}_h^R(\ulcorner\varphi\urcorner)$. A proof of $\text{PA} \vdash \sigma_1 \leftrightarrow \text{Pr}_h^R(\ulcorner\neg\varphi\urcorner)$ is similar, and we omit it. We argue in PA .

(\rightarrow): Suppose that σ_0 holds. Let n be the least number such that $\tau_0(n)$ holds. Then, either φ or φ^* holds, and hence either one of them is provable in T . Then, either φ or $\neg\varphi$ is T -provable. Let m be the least number such that $m \geq n$ and m is a T -proof of φ or $\neg\varphi$. Then, h does not output $\neg\varphi$ before Stage m , and $h(k_m) = \varphi$. This means $\text{Pr}_h^R(\ulcorner\varphi\urcorner)$ holds.

(\leftarrow): Suppose that $h(k_m) = \varphi$ and before Stage m , h does not output $\neg\varphi$. For the least such an m , $\exists z \leq m \tau_0(z)$ holds, and thus σ_0 holds. ■

COROLLARY 4.5. *For any Σ_1 sentence σ , the following are equivalent:*

1. *There exists a Σ_1 sentence σ' such that $\text{PA} \vdash \neg(\sigma \wedge \sigma')$ and $\text{PA} \vdash \neg\text{Con}_T \rightarrow \sigma \vee \sigma'$.*
2. *There exists a Rosser provability predicate $\text{Pr}_T^R(x)$ of T and an \mathcal{L}_A -sentence φ such that $\text{PA} \vdash \sigma \leftrightarrow \text{Pr}_T^R(\ulcorner\varphi\urcorner)$.*

PROOF. ($1 \Rightarrow 2$): Immediate from Theorem 4.2.

($2 \Rightarrow 1$): This implication is derived by letting σ' be the Σ_1 sentence $\text{Pr}_T^R(\ulcorner\neg\varphi\urcorner)$. ■

The $(\text{PA} + \text{Con}_T)$ -provable equivalence in the statement of the FGH theorem is equivalent to $\text{PA} \vdash (\sigma \vee \neg\text{Con}_T) \leftrightarrow \text{Pr}_T(\ulcorner\varphi\urcorner)$. In this viewpoint, the following corollary seems to be a natural counterpart of the FGH theorem in terms of Rosser provability predicates.

COROLLARY 4.6. *For any Σ_1 sentence σ , there exist a Rosser provability predicate $\text{Pr}_T^R(x)$ of T and an \mathcal{L}_A -sentence φ such that*

$$PA \vdash (\sigma \vee \neg \text{Con}_T) \leftrightarrow \text{Pr}_T^R(\ulcorner \varphi \urcorner).$$

PROOF. For the Σ_1 sentences $\sigma \vee \neg \text{Con}_T$ and $0 = 1$, we have that PA proves $\neg((\sigma \vee \neg \text{Con}_T) \wedge 0 = 1)$ and $\neg \text{Con}_T \rightarrow (\sigma \vee \neg \text{Con}_T) \vee 0 = 1$. By Corollary 4.5, there exist a Rosser provability predicate $\text{Pr}_T^R(x)$ of T and an \mathcal{L}_A -sentence φ such that $PA \vdash (\sigma \vee \neg \text{Con}_T) \leftrightarrow \text{Pr}_T^R(\ulcorner \varphi \urcorner)$. ■

Since $PA + \text{Con}_T \vdash \text{Pr}_T^R(\ulcorner \varphi \urcorner) \leftrightarrow \text{Pr}_T^R(\ulcorner \neg \varphi \urcorner)$, the FGH theorem directly follows from Corollary 4.6.

We show a version of the FGH theorem with respect to Rosser provability predicates corresponding to the representability of computable sets.

COROLLARY 4.7. *For any $\Delta_1(PA)$ sentence δ , there exist a Rosser provability predicate $\text{Pr}_T^R(x)$ of T and an \mathcal{L}_A -sentence φ such that*

$$PA \vdash \delta \leftrightarrow \text{Pr}_T^R(\ulcorner \varphi \urcorner) \text{ and } PA \vdash \neg \delta \leftrightarrow \text{Pr}_T^R(\ulcorner \neg \varphi \urcorner).$$

PROOF. Since δ is $\Delta_1(PA)$, there exist Σ_1 sentences σ_0 and σ_1 such that $PA \vdash \delta \leftrightarrow \sigma_0$ and $PA \vdash \neg \delta \leftrightarrow \sigma_1$. Then, we have $PA \vdash \neg(\sigma_0 \wedge \sigma_1)$ and $PA \vdash \sigma_0 \vee \sigma_1$. By Theorem 4.2, there exists a Rosser provability predicate $\text{Pr}_T^R(x)$ of T and an \mathcal{L}_A -sentence φ such that

$$PA \vdash \sigma_0 \leftrightarrow \text{Pr}_T^R(\ulcorner \varphi \urcorner) \text{ and } PA \vdash \sigma_1 \leftrightarrow \text{Pr}_T^R(\ulcorner \neg \varphi \urcorner).$$

■

Corollary 4.7 says that if a Σ_1 sentence σ is $\Delta_1(PA)$, then there exist a Rosser provability predicate $\text{Pr}_T^R(x)$ of T and an \mathcal{L}_A -sentence φ such that $PA \vdash \sigma \leftrightarrow \text{Pr}_T^R(\ulcorner \varphi \urcorner)$. Does this hold for all Σ_1 sentences? By Corollary 4.5, this question is rephrased as follows: For any Σ_1 sentence σ , does there exist a Σ_1 sentence σ' such that $PA \vdash \neg(\sigma \wedge \sigma')$ and $PA \vdash \neg \text{Con}_T \rightarrow \sigma \vee \sigma'$? We show that this is not the case.

PROPOSITION 4.8. *There exists a Σ_1 sentence σ such that for all Σ_1 sentences σ' , neither $PA \vdash \neg(\sigma \wedge \sigma')$ nor $PA \vdash \neg \text{Con}_T \rightarrow \sigma \vee \sigma'$. That is, for all Rosser provability predicates $\text{Pr}_T^R(x)$ of T and all \mathcal{L}_A -sentences φ , $PA \not\vdash \sigma \leftrightarrow \text{Pr}_T^R(\ulcorner \varphi \urcorner)$.*

PROOF. Let σ be a Σ_1 sentence which is Π_1 -conservative over PA such that $PA + \neg \text{Con}_T \not\vdash \sigma$. The existence of such a sentence is proved by Guaspari

[4, Theorem 2.4] (see also Lindström [11, Exercise 5.5 (b)]). Suppose that for some Σ_1 sentence σ' , $\text{PA} \vdash \neg(\sigma \wedge \sigma')$ and $\text{PA} \vdash \neg\text{Con}_T \rightarrow \sigma \vee \sigma'$. Then, $\text{PA} + \sigma \vdash \neg\sigma'$, and hence $\text{PA} \vdash \neg\sigma'$ by Π_1 -conservativity. Thus, $\text{PA} + \neg\text{Con}_T \vdash \sigma$. This is a contradiction. ■

Guaspari and Solovay [5] introduced the logic **R** of witness comparison formulas, and also Shavrukov [13] introduced the bimodal logic **GR** of usual and Rosser provability predicates. As in our observations in Section 3, it may also be possible to extend Theorem 4.2 to the framework of modal logic via these logics. For example, for **R**, we expect that the condition $\mathbf{R} + \{\Diamond^n \top \mid n \in \omega\} \not\vdash \Box\Box A \rightarrow \Box A$ works well.

Acknowledgements. This work was partly supported by JSPS KAKENHI Grant Number JP19K14586. The author would like to thank Sohei Iwata, Haruka Kogure, and Yuya Okawa for their helpful comments. The author would also like to thank the anonymous referees for their valuable comments and suggestions. In particular, the current proofs of Theorems 1 and 2 using Lemmas 1 and 2 are based on the ideas of one of the referees, which made the proofs significantly simpler and easier to understand.

Funding Open access funding provided by Kobe University.

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- [1] ARTEMOV, S. N., and L. D. BEKLEMISHEV, Provability logic, in D. Gabbay, and F. Guenther, (eds.), *Handbook of Philosophical Logic*, Springer, Dordrecht, 2005, pp. 189–360
- [2] BOLOS, G., *The logic of provability*, Cambridge University Press, Cambridge, 1993.

- [3] EHRENFEUCHT, A., and S. FEFERMAN, Representability of recursively enumerable sets in formal theories, *Archiv für Mathematische Logik und Grundlagenforschung*, 5:37–41, 1961.
- [4] GUASPARI, D., Partially conservative extensions of arithmetic, *Transactions of the American Mathematical Society*, 254:47–68, 1979.
- [5] GUASPARI, D., and R. M. SOLOVAY, Rosser sentences, *Annals of Mathematical Logic*, 16:81–99, 1979.
- [6] JAPARIDZE, G., and D. DE JONGH, The logic of provability, in S.R. Buss, (ed.), *Handbook of Proof Theory*, Elsevier, Amsterdam, 1998, pp. 475–546.
- [7] JOOSTEN, J. J., Turing jumps through provability, in A. Beckmann, V. Mitrana, and M. Soskova, (eds.), *Evolving Computability, 11th Conference on Computability in Europe, CiE 2015, Bucharest, Romania, June 29-July 3, 2015. Proceedings*, Springer, Cham, 2015, pp. 216–225.
- [8] KIKUCHI, M., and T. KURAHASHI, Illusory models of Peano arithmetic, *The Journal of Symbolic Logic*, 81(3):1163–1175, 2016.
- [9] KURAHASHI, T., Henkin sentences and local reflection principles for Rosser provability, *Annals of Pure and Applied Logic*, 167(2):73–94, 2016.
- [10] KURAHASHI, T., Provability logics relative to a fixed extension of Peano arithmetic, *The Journal of Symbolic Logic*, 83(3):1229–1246, 2018.
- [11] LINDSTRÖM, P., *Aspects of incompleteness*, 10. Natick, MA: Association for Symbolic Logic, 2003.
- [12] SEGERBERG, K., *An essay in classical modal logic*, 1, 2, 3. Filosofiska Studier, no. 13, University of Uppsala, Uppsala, 1971.
- [13] SHAVRUKOV, V. YU., On Rosser’s provability predicate, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 37(4):317–330, 1991.
- [14] SHEPHERDSON, J. C., Representability of recursively enumerable sets in formal theories, *Archiv für Mathematische Logik und Grundlagenforschung*, 5:119–127, 1961.
- [15] SMORYŃSKI, C., Fifty years of self-reference in arithmetic, *Notre Dame Journal of Formal Logic*, 22:357–374, 1981.
- [16] SOLOVAY, R. M., Provability interpretations of modal logic, *Israel Journal of Mathematics*, 25:287–304, 1976.
- [17] VISSER, A., The provability logics of recursively enumerable theories extending Peano arithmetic at arbitrary theories extending Peano arithmetic, *Journal of Philosophical Logic*, 13:97–113, 1984.
- [18] VISSER, A., Faith & falsity, *Annals of Pure and Applied Logic*, 131(1-3):103–131, 2005.

T. KURAHASHI
 Graduate School of System Informatics
 Kobe University
 1-1 Rokkodai Nada
 Kobe 657-8501
 Japan
kurahashi@people.kobe-u.ac.jp