The Σ_1-Provability Logic of HA^*

Mohammad Ardeshir[*]
Department of Mathematical Sciences
Sharif University of Technology
Tehran, Iran

Mojtaba Mojtahedi[†]
Department of Mathematics,
Statistics and Computer Science,
College of Sciences, University of Tehran

August 27, 2018

Abstract

For the Heyting Arithmetic HA, HA^* is defined [14, 15] as the theory $\{A \mid \mathsf{HA} \vdash A^\Box\}$, where $A^\Box$ is called the box translation of A (Definition 2.4). We characterize the Σ_1-provability logic of HA^* as a modal theory iH_σ^* (Definition 3.16).

Contents

1 Introduction

This paper is a sequel of our previous paper [2], in which we characterized the Σ_1-provability logic of HA as a decidable modal theory iH_σ (see Definition 3.16). Most of the materials of this paper are from the paper mentioned above. Our techniques and proofs are very similar to those used there. We use a crucial fact (Theorem 4.1 in this paper) proved in [2]. For the sake of self-containedness as much as possible, we bring here some definitions from that paper.

For an arithmetical theory T extending HA, the following axiom schema is called *the Completeness Principle*, CP_T:

$$A \to \Box_T A.$$

[*]mardeshir@sharif.ir
[†]http://mmojtahedi.ir/

Recall that by the work of Gödel in [5], for each arithmetical formula A and recursively axiomatizable theory T (like *Peano Arithmetic* PA), we can formalize the statement "there exists a proof in T for A" by a sentence of the language of arithmetic, i.e. $\mathsf{Prov}_T(\ulcorner A \urcorner) := \exists x\, \mathsf{Proof}_T(x, \ulcorner A \urcorner)$, where $\ulcorner A \urcorner$ is the code of A. Now, by *interpreting* $\Box_T$ by $\mathsf{Prov}_T(\ulcorner A \urcorner)$, the completeness principle for theory T is read as follows:

$$A \to \mathsf{Prov}_T(\ulcorner A \urcorner).$$

Albert Visser in [14, 15] introduced an extension of HA,

$$\mathsf{HA}^* := \mathsf{HA} + \mathsf{CP}_{\mathsf{HA}^*}.$$

He called HA^* as a *self-completion* of HA. Moreover, he showed that HA^* may be defined as the theory $\{A \mid \mathsf{HA} \vdash A^{\Box}\}$, where $A^{\Box}$ is called the *box translation* of A (Definition 2.4).

The notion of *provability logic* goes back essentially to K. Gödel [6] in 1933. He intended to provide a semantics for Heyting's formalization of *intuitionistic logic* IPC. He defined a *translation*, or *interpretation* τ from the propositional language to the modal language such that

$$\mathsf{IPC} \vdash A \qquad \Longleftrightarrow \qquad \mathsf{S4} \vdash \tau(A).$$

Now the question is whether we can find some modal propositional theory such that the $\Box$ operator captures the notion of *provability* in Peano Arithmetic PA. Hence the question is to find some propositional modal theory $T_{\Box}$ such that:

$$T_{\Box} \vdash A \qquad \Longleftrightarrow \qquad \forall * \ \mathsf{PA} \vdash A^*$$

By $(\ )^*$, we mean a mapping from the modal language to the first-order language of arithmetic, such that

- p^* is an arithmetical first-order sentence, for any atomic variable p, and $\bot^* = \bot$,
- $(A \circ B)^* = A^* \circ B^*$, for $\circ \in \{\vee, \wedge, \to\}$,
- $(\Box A)^* := \exists x\, \mathsf{Proof}_{\mathsf{PA}}(x, \ulcorner A^* \urcorner)$.

It turned out that $\mathsf{S4}$ is *not* a right candidate for interpreting the notion of *provability*, since $\neg \Box \bot$ is a theorem of $\mathsf{S4}$, contradicting Gödel's second incompleteness theorem (Peano Arthmetic PA, does not prove its own consistency).

Martin Löb in 1955 showed [10] that the Löb's rule $(\Box A \to A / A)$ is valid. Then in 1976, Robert Solovay [12] proved that the right modal logic, in which the $\Box$ operator interprets the notion of *provability in* PA, is GL. This modal logic is well-known as the Gödel-Löb logic, and has the following axioms and rules:

- all tautologies of classical propositional logic,
- $\Box(A \to B) \to (\Box A \to \Box B)$,
- $\Box A \to \Box\Box A$,
- Löb's axiom (L): $\Box(\Box A \to A) \to \Box A$,
- Necessitation Rule: $A / \Box A$,
- Modus ponens: $(A, A \to B)/B$.

Theorem. (Solovay-Löb) *For any sentence A in the language of modal logic, $\mathsf{GL} \vdash A$ if and only if for all interpretations $(\ )^*$, $\mathsf{PA} \vdash A^*$.*

Now let we restrict the map $(\ )^*$ on the atomic variables in the following sense. For any atomic variable p, $(p)^*$ is a Σ_1 sentence. This translation or interpretation is called Σ_1-interpretation. On the other hand, let $\mathsf{GLV} = \mathsf{GL} + \mathsf{CP}_a$, where CP_a is the completeness principle restricted to atomic variables, i.e., $p \to \Box p$. Albert Visser [14] proved the following result:

Theorem. (Visser) *For any sentence A in the language of modal logic, $\mathsf{GLV} \vdash A$ if and only if for all Σ_1 interpretations $(\ )^*$, $\mathsf{PA} \vdash A^*$.*

The question of generalizing Solovay's result from classical theories to intuitionistic ones, such as the intuitionistic counterpart of PA, well-known as HA, proved to be remarkably difficult and remains a major *open problem* since the end of 70s [3]. For a detailed history of the origins, backgrounds and motivations of the *provability logic*, we refer the readers to [3].

The following list contains crucial results about the provability logic of HA with arithmetical nature:

- John Myhill 1973 and Harvey Friedman 1975. $\mathsf{HA} \nvdash \Box_{\mathsf{HA}}(A \vee B) \to (\Box_{\mathsf{HA}} A \vee \Box_{\mathsf{HA}} B)$, [11, 4].
- Daniel Leivant 1975. $\mathsf{HA} \vdash \Box_{\mathsf{HA}}(A \vee B) \to \Box_{\mathsf{HA}}(\boxdot_{\mathsf{HA}} A \vee \boxdot_{\mathsf{HA}} B)$, in which $\boxdot_{\mathsf{HA}} A$ is a shorthand for $A \wedge \Box_{\mathsf{HA}} A$, [8].
- Albert Visser 1981. $\mathsf{HA} \vdash \Box_{\mathsf{HA}} \neg\neg \Box_{\mathsf{HA}} A \to \Box_{\mathsf{HA}} \Box_{\mathsf{HA}} A$ and $\mathsf{HA} \vdash \Box_{\mathsf{HA}}(\neg\neg \Box_{\mathsf{HA}} A \to \Box_{\mathsf{HA}} A) \to \Box_{\mathsf{HA}}(\Box_{\mathsf{HA}} A \vee \neg \Box_{\mathsf{HA}} A)$, [14, 15].
- Rosalie Iemhoff 2001 introduced a uniform axiomatization of all known axiom schemas of the provability logic of HA in an extended language with a bimodal operator $\rhd$. In her Ph.D. dissertation [7], Iemhoff raised a conjecture that implies directly that her axiom system, iPH, restricted to the normal modal language, is equal to the provability logic of HA, [7].
- Albert Visser 2002 introduced a decision algorithm for $\mathsf{HA} \vdash A$, for all modal propositions A not containing any atomic variable, i.e. A is made up of $\top, \bot$ via the unary modal connective $\Box_{\mathsf{HA}}$ and propositional connectives $\vee, \wedge, \to$, [16].
- Mohammad Ardeshir and Mojtaba Mojtahedi 2014 characterized the Σ_1-provability logic of HA as a decidable modal theory [2], named there and here as iH_σ. Recently, Albert Visser and Jetze Zoethout [18] proved this result by an alternative method.

The authors of [1] found a *reduction* of the Solovay-Löb Theorem to the Visser Theorem *only by propositional substitutions* [1]. This result is tempting to think of applying similar method for the intuitionistic case. However it seems to us that there is no obvious way of doing such reduction for the intuitionistic case, and it should be more complicated than the classical case.

In this paper, we introduce an axiomatization of a decidable modal theory iH_σ^* (see Definition 3.16) and prove that it is the Σ_1-provability logic of HA^*. This arithmetical theory is defined [14, 15] as the theory $\{A \mid \mathsf{HA} \vdash A^\Box\}$, where $A^\Box$ is called *the box translation* of A (Definition 2.4). It is worth mentioning that our proof of the Σ_1-provability logic of HA^* is in some sense, a *reduction* to the proof of the Σ_1-provability logic of HA, *only by propositional modal logic.*

2 Definitions, conventions and basic facts

The propositional non-modal language contains atomic variables, $\vee, \wedge, \to, \bot$ and propositional modal language is propositional non-modal language plus $\Box$. We use $\boxdot A$ as a shorthand for $A \wedge \Box A$. For simplicity, in this paper we use propositional language instead of propositional *modal* language. IPC is the intuitionistic propositional non-modal logic over usual propositional non-modal language. $\mathsf{IPC}_\Box$ is the same theory IPC in the extended language of propositional modal language, i.e. its language is propositional modal language and its axioms and rules are the same as the one in IPC. Since we have no axioms for $\Box$ in $\mathsf{IPC}_\Box$, it is obvious that $\Box A$ for each A, behave exactly like an atomic variable inside $\mathsf{IPC}_\Box$. Note that nothing more than symbol of A plays a role in $\Box A$. The first-order intuitionistic theory is denoted with IQC and CQC is its classical closure, i.e. IQC plus the principle of excluded middle. We have the usual first-order language of arithmetic which has a primitive recursive function symbol for each primitive recursive function. We use the same notations and definitions for Heyting's arithmetic HA as in [13], and Peano Arithmetic PA is HA plus the principle of excluded middle. For a set of sentences and rules $\Gamma \cup \{A\}$ in propositional non-modal, propositional modal or first-order language, $\Gamma \vdash A$ means that A is derivable from Γ in the system $\mathsf{IPC}, \mathsf{IPC}_\Box, \mathsf{IQC}$, respectively.

Definition 2.1. *Suppose T is an r.e arithmetical theory and σ is a function from atomic variables to arithmetical sentences. We extend σ to all modal propositions A, inductively:*

- $\sigma_T(A) := \sigma(A)$ *for atomic A,*
- σ_T *distributes over $\wedge, \vee, \to$,*
- $\sigma_T(\Box A) := Pr_T(\ulcorner \sigma_T(A) \urcorner)$, *in which $Pr_T(x)$ is the Σ_1-predicate that formalizes provability of a sentence with Gödel number x, in the theory T.*

We call σ to be a Σ_1-substitution, if for every atomic A, $\sigma(A)$ be a Σ_1-formula.

Definition 2.2. *Provability logic of a sufficiently strong theory, T is defined to be a modal propositional theory $\mathcal{PL}(T)$ such that $\mathcal{PL}(T) \vdash A$ iff for arbitrary arithmetical substitution σ_T, $T \vdash \sigma_T(A)$. If we restrict the substitutions to Σ-substitutions, then the new modal theory is $\mathcal{PL}_\sigma(T)$.*

Lemma 2.3. *Let $A(p_1, \ldots, p_n)$ be a non-modal proposition with $p_i \neq p_j$ for all $0 < i < j \leq n$. Then for every modal sentences $B_1, \ldots, B_n$ with $B_i \neq B_j$ for $0 < i < j \leq n$ we have:*

$$\mathsf{IPC} \vdash A \quad \textit{iff} \quad \mathsf{IPC}_\Box \vdash A[p_1|\Box B_1, \ldots, p_n|\Box B_n].$$

Proof. By simple inductions on complexity of proofs in IPC and $\mathsf{IPC}_\Box$. ❏

The following definition, the Beeson-Visser box-translation, is essentially from ([15, Def.4.1]).

Definition 2.4. *For every proposition A in modal propositional language, we associate a proposition $A^\Box$, called box-translation of A, defined inductively as follows:*

- $A^\Box := A \wedge \Box A$, *for atomic A, and $\bot^\Box = \bot$,*
- $(A \circ B)^\Box := A^\Box \circ B^\Box$, *for $\circ \in \{\vee, \wedge\}$,*
- $(A \to B)^\Box := (A^\Box \to B^\Box) \wedge \Box(A^\Box \to B^\Box)$,
- $(\Box A)^\Box := \Box(A^\Box)$.

The box-translation can be extended to first-order arithmetical formulae A, as follows:

- $(\forall x A)^\Box := \Box(\forall x A^\Box) \wedge \forall x A^\Box$,
- $(\exists x A)^\Box := \exists x A^\Box$.

Define NOI (No Outside Implication) as set of modal propositions A, that any occurrence of $\to$ is in the scope of some $\Box$. To be able to state an extension of Leivant's Principle (that is adequate to axiomatize Σ_1-provability logic of HA) we need a translation on modal language which we name it Leivant's translation. We define it recursively as follows:

- $A^l := A$ for atomic A, boxed A or $A = \bot$,

- $(A \wedge B)^l := A^l \wedge B^l$,
- $(A \vee B)^l := \boxdot A^l \vee \boxdot B^l$,
- $(A \to B)^l$ is defined by cases: If $A \in \mathsf{NOI}$, define $(A \to B)^l := A \to B^l$, else define $(A \to B)^l := A \to B$.

Definition 2.5. *Minimal provability logic $i\mathsf{GL}$, is same as Gödel-Löb provability logic GL , without the principle of excluded middle, i.e. it has the following axioms and rules:*

- *theorem of* $\mathsf{IPC}_\Box$,
- $\Box(A \to B) \to (\Box A \to \Box B)$,
- $\Box A \to \Box\Box A$,
- *Löb's axiom* (L)*:* $\Box(\Box A \to A) \to \Box A$,
- *Necessitation Rule:* $A/\Box A$,
- *Modus ponens:* $(A, A \to B)/B$.

2.1 Definition of modal theories

$i\boldsymbol{K}4$ is $i\mathsf{GL}$ without Löb's axiom. Note that we can get rid of the necessitation rule by adding $\Box A$ to the axioms, for each axiom A in the above list. We will use this fact later in this paper. We list the following axiom schemae:

- *The Completeness Principle:* $\mathsf{CP} := A \to \Box A$.
- *Restricted Completeness Principle to atomic formulae:* $\mathsf{CP}_a := p \to \Box p$, *for atomic* p.
- *Leivant's Principle:* $\mathsf{Le} := \Box(B \vee C) \to \Box(\Box B \vee C)$. *[9]*
- *Extended Leivant's Principle:* $\mathsf{Le}^+ := \Box A \to \Box A^l$.
- *Trace Principle:* $\mathsf{TP} := \Box(A \to B) \to (A \vee (A \to B))$. *[15]*

We define theories $\mathsf{iGLC} := i\mathsf{GL} + \mathsf{CP}$, $\mathsf{H} := \mathsf{iGLC} + \mathsf{TP}$, $\mathsf{LLe} := \mathsf{iGL} + \mathsf{Le}$ *and* $\mathsf{LLe}^+ := i\mathsf{GL} + \mathsf{Le}^+ + \mathsf{CP}_a$. *Note that in the presence of* CP *and modus ponens, the necessitation rule is superfluous. Later we will find Kripke semantics for* iGLC *and also we will see that* iGLC *and* LLe^+ *proves the same formulae of restricted complexity* (TNNIL).

2.2 HA^* and PA^*

HA^* and PA^* were first introduced in [15]. These theories are defined as

$$\mathsf{HA}^* := \{A \mid \mathsf{HA} \vdash A^\Box\} \quad \text{and} \quad \mathsf{PA}^* := \{A \mid \mathsf{PA} \vdash A^\Box\}.$$

Visser in [15] showed that the provability logic of PA^* is H, i.e. $\mathsf{H} \vdash A$ iff for all arithmetical substitution σ, $\mathsf{PA}^* \vdash \sigma_{\mathsf{PA}^*}(A)$. That means that

$$\mathcal{PL}(\mathsf{PA}^*) = \mathcal{PL}_\sigma(\mathsf{PA}^*) = \mathsf{H}.$$

Lemma 2.6. *1. For any arithmetical Σ_1-formula A,* $\mathsf{HA} \vdash A \leftrightarrow A^\Box$.

2. HA *is closed under the box-translation, i.e., for any arithmetical formula A,* $\mathsf{HA} \vdash A$ *implies* $\mathsf{HA} \vdash A^\Box$, *so* $\mathsf{HA} \subseteq \mathsf{HA}^*$.

Proof. 1. See [15](4.6.iii).

2. See [15](4.14.i).

❑

Lemma 2.7. *For any Σ_1-substitution σ and each propositional modal sentence A, we have* $\mathsf{HA} \vdash \sigma_{\mathsf{HA}}(A^{\Box}) \leftrightarrow (\sigma_{\mathsf{HA}^*}(A))^{\Box}$ *and hence*

$$\mathsf{HA} \vdash \sigma_{\mathsf{HA}}(A^{\Box}) \quad \textit{iff} \quad \mathsf{HA}^* \vdash \sigma_{\mathsf{HA}^*}(A)$$

Proof. Use induction on the complexity of A. All the steps are straightforward. For the atomic case, we use Lemma 2.6.1. ❑

Remark 2.8. This lemma can be combined with the characterization of the Σ_1-provability logic of HA to derive directly a characterization of the Σ_1-provability logic of HA^*:

A belongs to the Σ_1-provability logic of HA^* iff $A^{\Box}$ belongs to the Σ_1-provability logic of HA.

This means that we have a decision algorithm for the Σ_1-provability logic of HA^*. The rest of this paper is devoted to axiomatize the Σ_1-provability logic of HA^*.

3 Propositional modal logics

3.1 NNIL formulae and related topics

The class of *No Nested Implications in the Left*, NNIL formulae in a propositional language was introduced in [17] , and more explored in [16]. The crucial result of [16] is providing an algorithm that as input, gets a non-modal proposition A and returns its best NNIL approximation A^* from below, i.e., $\mathsf{IPC} \vdash A^* \to A$ and for all NNIL formula B such that $\mathsf{IPC} \vdash B \to A$, we have $\mathsf{IPC} \vdash B \to A^*$. In the following we explain this algorithm and explain how to extend it to modal propositions.

To define the class of NNIL propositions, let us first define a complexity measure ρ on non-modal propositions as follows:

- $\rho p = \rho\bot = \rho\top = 0$, where p is an atomic proposition,
- $\rho(A \wedge B) = \rho(A \vee B) = max(\rho A, \rho B)$,
- $\rho(A \to B) = max(\rho A + 1, \rho B)$,

Then $\mathsf{NNIL} = \{A \mid \rho A \leq 1\}$.

Definition 3.1. *We define a measure complexity for modal propositions D as follows:*

- $I(D) := \{E \in Sub(D) \mid E$ *is an implication that is not in the scope of a* $\Box\}$,
- $\mathfrak{i}(D) := max\{|I(E)| \mid E \in I(D)\}$, *where $|X|$ is the number of elements of X,*
- $\mathfrak{c}D :=$ *the number of occurrences of logical connectives which is not in the scope of a* $\Box$,
- $\mathfrak{d}D :=$ *the maximum number of nested boxes. To be more precise,*
  - $\mathfrak{d}D := 0$ *for atomic D,*
  - $\mathfrak{d}D := max\{\mathfrak{d}D_1, \mathfrak{d}D_2\}$, *where $D = D_1 \circ D_2$ and $\circ \in \{\wedge, \vee, \to\}$,*
  - $\mathfrak{d}\Box D := \mathfrak{d}D + 1$,
- $\mathfrak{o}D := (\mathfrak{d}D, \mathfrak{i}D, \mathfrak{c}D)$.

Note that the measure $\mathfrak{o}D$ *is ordered lexicographically, i.e.,* $(d, i, c) < (d', i', c')$ *iff* $d < d'$ *or* $d = d', i < i'$ *or* $d = d', i = i', c < c'$.

Definition 3.2. *For any two modal propositions* A, B*, we define* $[A]B$ *and* $[A]'B$*, by induction on the complexity of* B*:*

- $[A]p = [A]'p = p$*, for atomic* p*,* $\top$ *and* $\bot$*,*
- $[A](B_1 \circ B_2) = [A](B_1) \circ [A](B_2)$*,* $[A]'(B_1 \circ B_2) = [A]'(B_1) \circ [A]'(B_2)$ *for* $\circ \in \{\vee, \wedge\}$*,*
- $[A](B_1 \to B_2) = A \to (B_1 \to B_2)$*,* $[A]'(B_1 \to B_2) = (A' \wedge B_1) \to B_2$*, in which* $A' = A[B_1 \to B_2 \mid B_2]$*, i.e., replace each occurrence of* $B_1 \to B_2$ *in* A *by* B_2*,*

NNIL-algorithm

For each proposition A, A^* is produced by induction on complexity measure $\mathfrak{o}A$ as follows. For details see [16].

1. A is atomic, take $A^* := A$,
2. $A = B \wedge C$, take $A^* := B^* \wedge C^*$,
3. $A = B \vee C$, take $A^* := B^* \vee C^*$,
4. $A = B \to C$, we have several sub-cases. In the following, an occurrence of E in D is called an *outer occurrence*, if E is not in the scope of an implication.
   4.a. C contains an outer occurrence of a conjunction. In this case, we assume some formula $J(q)$ such that
   - q is a propositional variable not occurred in A,
   - q is outer in J and occurs exactly once,
   - $C = J[q|(D \wedge E)]$.

   Such J obviously exists. Now set $C_1 := J[q|D], C_2 := J[q|E]$ and $A_1 := B \to C_1, A_2 := B \to C_2$ and finally, define $A^* := A_1^* \wedge A_2^*$.
   4.b. B contains an outer occurrence of a disjunction. In this case we suppose some formula $J(q)$ such that
   - q is a propositional variable not occurred in A,
   - q is outer in J and occurs exactly once,
   - $B = J[q|(D \vee E)]$.

   Such J obviously exists. Now set $B_1 := J[q|D], B_2 := J[q|E]$ and $A_1 := B_1 \to C, A_2 := B_2 \to C$ and finally, define $A^* := A_1^* \wedge A_2^*$.

   4.c. $B = \bigwedge X$ and $C = \bigvee Y$ and X, Y are sets of implications or atoms. We have several sub-cases:

   4.c.i. X contains atomic p. Set $D := \bigwedge(X \setminus \{p\})$ and take $A^* := p \to (D \to C)^*$.

   4.c.ii. X contains $\top$. Define $D := \bigwedge(X \setminus \{\top\})$ and take $A^* := (D \to C)^*$.

   4.c.iii. X contains $\bot$. Take $A^* := \top$.

   4.c.iv. X contains only implications. For any $D = E \to F \in X$, let

$$B \downarrow D := \bigwedge((X \setminus \{D\}) \cup \{F\}).$$

Let $Z := \{E \mid E \to F \in X\} \cup \{C\}$ and $A_0 := [B]Z := \bigvee\{[B]E \mid E \in Z\}$. Now if $\mathfrak{o}A_0 < \mathfrak{o}A$, we take

$$A^* := \bigwedge\{((B \downarrow D) \to C)^* | D \in X\} \wedge A_0^*,$$

otherwise, first set $A_1 := [B]'Z$ and then take

$$A^* := \bigwedge\{((B \downarrow D) \to C)^* | D \in X\} \wedge A_1^*$$

We can extend ρ to all modal language with $\rho(\Box A) := 0$. The class of NNIL propositions may be defined for propositional modal language as well, i.e. we call a modal proposition A to be $\mathsf{NNIL}_\Box$, if $\rho(A) \leq 1$ (for extended ρ). We also define two other classes of propositions:

Definition 3.3. TNNIL (*Thoroughly* NNIL) *is the smallest class of propositions such that*

- TNNIL *contains all atomic propositions,*
- *if* $A, B \in \mathsf{TNNIL}$, *then* $A \vee B, A \wedge B, \Box A \in \mathsf{TNNIL}$,
- *if all* $\to$ *occurred in* A *are contained in the scope of a* $\Box$ *(or equivalently* $A \in \mathsf{NOI}$*) and* $A, B \in \mathsf{TNNIL}$, *then* $A \to B \in \mathsf{TNNIL}$.

Finally we define TNNIL^- *as the set of all the propositions like* $A(\Box B_1, \ldots, \Box B_n)$, *such that* $A(p_1, \ldots, p_n)$ *is an arbitrary non-modal proposition and* $B_1, \ldots, B_n \in \mathsf{TNNIL}$.

We can use the same algorithm with slight modifications treating propositions inside $\Box$ as well. First we extend Definition 3.2 to capture modal language.

Definition 3.4. *For any two modal propositions* A, B, *we define* $[A]B$ *and* $[A]'B$ *by induction on the complexity of* B. *We extend Definition 3.1 by the following item:*

- $[A]\Box B_1 = [A]'\Box B_1 := \Box B_1$.

It is clear that we are treating a boxed formula as an atomic variable.

$\mathsf{NNIL}_\Box$-algorithm

We use NNIL-algorithm with the following changes to produce a similar NNIL-algorithm for modal language.

1. A is atomic or boxed, take $A^* = A$.

4. An occurrence of E in D is called an *outer occurrence*, if E is neither in the scope of an implication nor in the scope of a boxed formula.

4. c(i). X contains atomic or boxed formula p. We set $D := \bigwedge(X \setminus \{p\})$ and take $A^* := p^* \to (D \to C)^*$.

Remark 3.5. In fact, we have two ways to find out $\mathsf{NNIL}_\Box$ approximation of a modal proposition.

First: simply apply $\mathsf{NNIL}_\Box$-algorithm to a modal proposition A and compute A^*.

Second: let $B_1, \ldots, B_n$ be all boxed sub-formulae of A which are not in the scope of any other boxes. Let $A'(p_1, \ldots, p_n)$ be unique non-modal proposition such that $\{p_i\}_{1 \leq i \leq n}$ are fresh atomic variables not occurred in A and $A = A'[p_1|B_1, \ldots, p_n|B_n]$. Let $\rho(A) := (A')^*[p_1|B_1, \ldots, p_n|B_n]$. Then it is easy to observe that $\mathsf{IPC}_\Box \vdash \rho(A) \leftrightarrow A^*$.

The above defined algorithm is not deterministic, however by the following Theorem, we know that A^* is unique up to $\mathsf{IPC}_\Box$ equivalence. The notation $A \rhd_{\mathsf{IPC}_\Box, \mathsf{NNIL}_\Box} B$ (A, $\mathsf{NNIL}_\Box$-preserves B) from [16], means that for each $\mathsf{NNIL}_\Box$ modal proposition C, if $\mathsf{IPC}_\Box \vdash C \to A$, then $\mathsf{IPC}_\Box \vdash C \to B$, in which A, B are modal propositions.

Theorem 3.6. *For each modal proposition* A, $\mathsf{NNIL}_\Box$ *algorithm with input* A *terminates and the output formula* A^*, *is an* $\mathsf{NNIL}_\Box$ *proposition such that* $\mathsf{IPC}_\Box \vdash A^* \to A$.

Proof. See [2, The. 4.5]. ❑

TNNIL-algorithm

Here we define A^+ as TNNIL-formula approximating A. Informally speaking, to find A^+, we first compute A^* and then replace all outer boxed formula $\Box B$ in A by $\Box B^+$. To be more accurate, we define A^+ by induction on $\mathfrak{d}A$. Suppose that for all B with $\mathfrak{d}B < \mathfrak{d}A$, we have defined B^+. Suppose that $A'(p_1, \ldots, p_n)$ and $\Box B_1, \ldots, \Box B_n$ such that $A = A'[p_1|\Box B_1, \ldots, p_n|\Box B_n]$ where A' is a non-modal proposition and $p_1, \ldots, p_n$ are fresh atomic variables (not occurred in A). It is clear that $\mathfrak{d}B_i < \mathfrak{d}A$ and then we can define $A^+ := (A')^*[p_1|\Box B_1^+, \ldots, p_n|\Box B_n^+]$.

Lemma 3.7. *For any modal proposition A,*

1. *for all Σ_1-substitution σ we have $\mathsf{HA} \vdash \Box\sigma_{\mathsf{HA}}(A) \leftrightarrow \Box\sigma_{\mathsf{HA}}(A^+)$ and hence $\mathsf{HA} \vdash \sigma_{\mathsf{HA}}(A)$ iff $\mathsf{HA} \vdash \sigma_{\mathsf{HA}}(A^+)$.*
2. *$i\mathsf{GL} \vdash A_1 \to A_2$ implies $i\mathsf{GL} \vdash A_1^+ \to A_2^+$, and $i\mathsf{K4} \vdash A_1 \to A_2$ implies $i\mathsf{K4} \vdash A_1^+ \to A_2^+$.*
3. *$i\mathsf{GL} \vdash A_1 \leftrightarrow A_2$ implies $i\mathsf{GL} \vdash A_1^+ \leftrightarrow A_2^+$, and $i\mathsf{K4} \vdash A_1 \leftrightarrow A_2$ implies $i\mathsf{K4} \vdash A_1^+ \leftrightarrow A_2^+$.*

Proof. See [2, Corollary 4.8]. ❑

TNNIL$^\Box$-algorithm

Corollary 3.8. *There exists a TNNIL$^\Box$-algorithm such that for any modal proposition A, it halts and produces a proposition $A^- \in \mathsf{TNNIL}^\Box$ such that $\mathsf{IPC}_\Box \vdash A^+ \to A^-$.*

Proof. Let $A := B(\Box C_1, \ldots, \Box C_n)$, and $B(p_1, \ldots, p_n)$ is non-modal. apparently such B exists. Then define $A^- := B(\Box C_1^+, \ldots, \Box C_n^+)$. Now definition of A^+ implies $A^+ = (A^-)^*$ and hence Theorem 3.6 implies that A^- has desired property. ❑

Lemma 3.9. *For each modal proposition A and Σ_1-substitution σ, $\mathsf{HA} \vdash \sigma_{\mathsf{HA}}A \leftrightarrow \sigma_{\mathsf{HA}}A^-$.*

Proof. Use definition of $(.)^-$ and Lemma 3.7.1. ❑

Remark 3.10. Note that $\mathsf{iGLC} \vdash A \leftrightarrow B$ does not imply $\mathsf{iGLC} \vdash A^+ \leftrightarrow B^+$. A counter-example is $A := \neg\neg p$ and $B := \neg \boxdot (\neg p)$. We have $A^+ = A^* = p$ and $\mathsf{iGLC} \vdash B^+ \leftrightarrow (\Box\neg p \to p)$. Now one can use Kripke models to show $\mathsf{iGLC} \nvdash \neg\neg p \to (\Box\neg p \to p)$.

Remark 3.11. In the $\mathsf{NNIL}_\Box$-algorithm, if we replace the operation $(\cdot)^*$ by $(\cdot)^\dagger$, and change the step 1 to
1. $A^\dagger := A$, if A is atomic, and $(\Box B)^\dagger := \Box B^\dagger$,
then the new algorithm also halts, and for any modal proposition A, we have $i\mathsf{K4} \vdash A^\dagger \leftrightarrow A^+$.

3.2 Box translation and propositional theories

Following Visser's definition of the notion of a *base* in arithmetical theories [15], we define

Definition 3.12. *A modal theory T is called to be* closed under box-translation *if for every proposition A, $T \vdash A$ implies $T \vdash A^\Box$.*

Proposition 3.13. *For arbitrary subset X of $\{\mathsf{CP}, \mathsf{CP}_a, \mathsf{L}\}$, $i\mathsf{K4}+X$ is closed under box-translation.*

Proof. The proof can be carried out in three steps:

1. For any proposition A first we show that $\mathsf{IPC}_\Box \vdash A$ implies $i\mathsf{K4} \vdash A^\Box$. This can be done by a routine induction on the length of the proof in IPC. Note that for any axiom A of IPC, we have $i\mathsf{K4} \vdash A^\Box$. As for the rule of modus ponens, suppose that $\mathsf{IPC}_\Box \vdash A$ and $\mathsf{IPC}_\Box \vdash A \to B$. By induction hypothesis, then $i\mathsf{K4} \vdash A^\Box$ and $i\mathsf{K4} \vdash (A^\Box \to B^\Box) \wedge \Box(A^\Box \to B^\Box)$ and so $i\mathsf{K4} \vdash B^\Box$.

2. Next observe that

$$(\Box A \to \Box\Box A)^{\Box} = \Box A^{\Box} \to \Box\Box A^{\Box}$$

and also

$$i\mathsf{K4} \vdash [(\Box(A \to B) \wedge \Box A) \to \Box B]^{\Box} \leftrightarrow (\Box(A^{\Box} \to B^{\Box}) \wedge \Box A^{\Box}) \to \Box B^{\Box}.$$

3. We observe that for any axiom $A \in X$, $i\mathsf{K4} + X \vdash A^{\Box}$.

❑

The following two lemmas will be used in the proof of Theorem 3.18.

Lemma 3.14. *For any modal propositions A, A' and B, and any propositional modal theory T containing $\mathsf{IPC}_{\Box}$,*

1. $i\mathsf{K4} + \Box A^{\Box} \vdash ([A]B)^{\Box} \leftrightarrow ([A^{\Box}]B^{\Box})$.
2. $T \vdash A \leftrightarrow A'$ *implies* $T \vdash [A]B \leftrightarrow [A']B$.

Proof. Proof of both parts are by induction on the complexity of B:

1. The only non-trivial case is when B is an implication. Let $B := C \to D$. By Definitions 3.4 and 2.4,

$$([A](C \to D))^{\Box} = \boxdot(A^{\Box} \to ((C^{\Box} \to D^{\Box}) \wedge \Box(C^{\Box} \to D^{\Box})))$$

and also

$$[A^{\Box}](C \to D)^{\Box} = (A^{\Box} \to (C^{\Box} \to D^{\Box})) \wedge \Box(C^{\Box} \to D^{\Box}).$$

Now it is easy to observe that

$$i\mathsf{K4} + \Box A^{\Box} \vdash ([A](C \to D))^{\Box} \leftrightarrow ([A^{\Box}](C \to D)^{\Box}).$$

2. Similar to the first item.

❑

Notation. In the sequel of paper, we use $A \equiv B$ as a shorthand for $i\mathsf{K4} \vdash A \leftrightarrow B$.

Lemma 3.15. *Let $A = B \to C$ be a modal proposition such that $B = \bigwedge X$ and $C = \bigvee Y$, where X is a set of implications and Y is a set of atomic, boxed or implicative propositions. Then*

$$(A^{\Box})^{+} \equiv \boxdot\left(\bigwedge_{E\to F\in X} \Box\left((E \to F)^{\Box}\right)^{+} \to \left(\bigwedge\left\{\left((B\downarrow D \to C)^{\Box}\right)^{+} \mid D \in X\right\} \wedge \left(([B]Z)^{\Box}\right)^{+}\right)\right)$$

where $Z = \{E | E \to F \in X\} \cup \{C\}$.

Proof. To simplify notations, Let us indicate

- the sets of all atomic and boxed propositions by At and Bo, respectively,
- $X' := \{E^{\Box} \to F^{\Box} \mid E \to F \in X\}$,
- $Z' := Z^{\Box} = \{E^{\Box} \mid E \to F \in X\} \cup \{C^{\Box}\}$,
- $B' := \bigwedge X'$,
- for any $I \subseteq Y$, $C^{I} := \bigvee_{E\to F\in I} \Box(E^{\Box} \to F^{\Box}) \vee \bigvee_{E\in I\cap\mathsf{At}} \Box E \vee \bigvee_{E\to F\in Y\setminus I}(E^{\Box} \to F^{\Box})$
$\vee \bigvee_{E\in(Y\setminus I)\cap\mathsf{At}} E \vee \bigvee_{E\in\mathsf{Bo}\cap Y} E^{\Box}$,
- and $Z^{I} := \{E^{\Box} \mid E \to F \in X\} \cup \{C^{I}\}$.

By repeated use of distributivity of conjunction over disjunction, which is valid in IPC , we have

$$C^{\Box} \equiv \bigwedge_{I\subseteq Y} C^I \quad \text{and} \quad Z^{\Box} \equiv \bigwedge_{I\subseteq Y} Z^I \tag{3.1}$$

Note that $A^{\Box} = (B^{\Box} \to C^{\Box}) \wedge \Box(B^{\Box} \to C^{\Box})$, and then by definition of $(\cdot)^+$,

$$(A^{\Box})^+ = (B^{\Box} \to C^{\Box})^+ \wedge \Box(B^{\Box} \to C^{\Box})^+.$$

Now we compute the left conjunct:

$$\left(B^{\Box} \to C^{\Box}\right)^+ = \bigwedge_{I\subseteq Y} \left(B^{\Box} \to C^I\right)^+ \tag{3.2}$$

$$\equiv \bigwedge_{I\subseteq Y} \left(\bigwedge_{E\to F\in X} \Box\left(E^{\Box} \to F^{\Box}\right)^+ \to \left(\left(\bigwedge_{E\to F\in X} \left(E^{\Box} \to F^{\Box}\right)\right) \to C^I\right)^+\right) \tag{3.3}$$

$$\equiv \bigwedge_{E\to F\in X} \Box\left((E \to F)^{\Box}\right)^+ \to \bigwedge_{I\subseteq Y} (B' \to C^I)^+ \tag{3.4}$$

$$\equiv \bigwedge_{E\to F\in X} \Box\left((E \to F)^{\Box}\right)^+ \to \bigwedge_{I\subseteq Y} \left(\bigwedge\left\{(B' \downarrow D' \to C^I)^+ \mid D' \in X'\right\} \wedge ([B']Z^I)^+\right) \tag{3.5}$$

$$\equiv \bigwedge_{E\to F\in X} \Box\left((E \to F)^{\Box}\right)^+ \to \left(\bigwedge\left\{\left(B' \downarrow D' \to C^{\Box}\right)^+ \mid D' \in X'\right\} \wedge ([B']Z')^+\right) \tag{3.6}$$

and hence

$$\left(A^{\Box}\right)^+ \equiv \boxdot\left(\bigwedge_{E\to F\in X} \Box\left((E \to F)^{\Box}\right)^+ \to \left(\bigwedge\left\{\left((B \downarrow D \to C)^{\Box}\right)^+ \mid D \in X\right\} \wedge ([B']Z')^+\right)\right) \tag{3.7}$$

Note that 3.2 and 3.3 hold by $\mathsf{NNIL}_{\Box}$-algorithm, 3.4 holds by properties of $\mathsf{IPC}_{\Box}$, 3.5 holds by TNNIL-algorithm, 3.6 holds by TNNIL-algorithm and equation 3.1, and finally equation 3.7 is derived from 3.6 by deduction in iK4 and TNNIL-algorithm. Now it is enough to show that the last formula is equivalent to the following one in iK4:

$$\boxdot\left(\bigwedge_{E\to F\in X} \Box\left((E \to F)^{\Box}\right)^+ \to \left(\bigwedge\left\{\left((B \downarrow D \to C)^{\Box}\right)^+ \mid D \in X\right\} \wedge \left(([B]Z)^{\Box}\right)^+\right)\right) \tag{3.8}$$

To show this, it is enough to show

$$i\mathsf{K4} \vdash \bigwedge_{E\to F\in X} \Box\left((E \to F)^{\Box}\right)^+ \to \left(\left(([B]Z)^{\Box}\right)^+ \leftrightarrow ([B']Z')^+\right).$$

Then by Lemma 3.7.2, it is enough to show $i\mathsf{K4} \vdash \bigwedge_{E\to F\in X} \Box(E \to F)^{\Box} \to (([B]Z)^{\Box} \leftrightarrow [B']Z')$. Since $\bigwedge_{E\to F\in X} \Box(E \to F)^{\Box} \equiv \Box B^{\Box}$, then it is enough to show $i\mathsf{K4} + \Box B^{\Box} \vdash ([B]Z)^{\Box} \leftrightarrow [B']Z'$. Now, by Lemma 3.14.1, we have $i\mathsf{K4} + \Box B^{\Box} \vdash ([B]Z)^{\Box} \leftrightarrow [B^{\Box}]Z^{\Box}$. Hence we should show $i\mathsf{K4} + \Box B^{\Box} \vdash [B^{\Box}]Z^{\Box} \leftrightarrow [B']Z'$. We have $Z' = Z^{\Box}$ and $i\mathsf{K4} + \Box B^{\Box} \vdash B^{\Box} \leftrightarrow B'$. Then by Lemma 3.14.2, $i\mathsf{K4} + \Box B^{\Box} \vdash [B^{\Box}]Z^{\Box} \leftrightarrow [B']Z'$. ❑

3.3 Axiomatizing TNNIL-algorithm

In this section, we introduce the axiom set X such that $i\mathsf{K4} + X \vdash (A^{\Box})^- \leftrightarrow A^{\Box}$. Note that we may simply choose $X := \{(A^{\Box})^- \leftrightarrow A^{\Box} \mid A \text{ is arbitrary proposition}\}$. However, we want to reduce X to some smaller efficient set of formulae.

We use some modal variant of Visser's $\rhd_\sigma$ in [16]. It is exactly the same as the relation $\rhd$ in [2] (sec. 4.3) except for item B2, which is a little bit different:

- B2$'$. Let X be a set of implications, $B := \bigwedge X$ and $A := B \to C$. Also assume that $Z := \{E|E \to F \in X\} \cup \{C\}$. Then $A \rhd [B]Z$,

The relation $\rhd^*$ is defined to be the smallest relation on modal propositional sentences satisfying:

- A1. If $i\mathsf{K4} \vdash A \to B$, then $A \rhd^* B$,
- A2. If $A \rhd^* B$ and $B \rhd^* C$, then $A \rhd^* C$,
- A3. If $C \rhd^* A$ and $C \rhd^* B$, then $C \rhd^* A \wedge B$,
- A4. If $A \rhd^* B$, then $\Box A \rhd^* \Box B$,
- B1. If $A \rhd^* C$ and $B \rhd^* C$, then $A \vee B \rhd^* C$,
- B2. Let X be a set of implications, $B := \bigwedge X$ and $A := B \to C$. Also assume that $Z := \{E|E \to F \in X\} \cup \{C\}$. Then $A \wedge \Box B \rhd^* [B]Z$,
- B3. If $A \rhd^* B$, then $p \to A \rhd^* p \to B$, in which p is atomic or boxed.

$A \rhd\lhd^* B$ means $A \rhd^* B$ and $B \rhd^* A$.

Definition 3.16. *We define*

$$\mathsf{iH}^*_\sigma := i\mathsf{GL} + \mathsf{CP} + \{\Box A \to \Box B | A \rhd^* B\}.$$

Note that the Σ_1-provability logic of HA is proved in [2] to be

$$\mathsf{iH}_\sigma := \mathsf{iGL} + \mathsf{CP_a} + \mathsf{Le}^+ + \{\Box \mathsf{A} \to \Box \mathsf{B} | \mathsf{A} \rhd \mathsf{B}\},$$

in which CP_a is the Completeness Principle restricted to atomic propositions.

Lemma 3.17. *For any propositional modal sentences A, B, $A \rhd^* B$ implies $A^\Box \rhd^* B^\Box$.*

Proof. It is clear that $A \rhd^* B$ iff there exists a Hilbert-type sequence of relations $\{A_i \rhd^* B_i\}_{0 \leq i \leq n}$ such that $A_n = A, B_n = B$ and for each $i \leq n$, $A_i \rhd^* B_i$ is an instance of axioms A1 or B2, or it is derived by making use of some previous members of sequence and some of the rules A2-A4 or B1 or B3. Hence we are authorized to use induction on the length of such sequence for $A \rhd^* B$ to show $A^\Box \rhd^* B^\Box$. The only non-trivial steps are axioms A1 and B2. Suppose that $A \rhd^* B$ is an instance of A1, i.e. $i\mathsf{K4} \vdash A \to B$. Then by Proposition 3.13, we have $i\mathsf{K4} \vdash A^\Box \to B^\Box$ and hence again by A1, $A^\Box \rhd^* B^\Box$, as desired.
For treating B2, suppose that $A := B \to C$, $B = \bigwedge X$, X is a set of implications and $Z := \{E|E \to F \in X\} \cup \{C\}$. We must show $(A \wedge \Box B)^\Box \rhd^* ([B]Z)^\Box$. Define $X' := \{E^\Box \to F^\Box | E \to F \in X\}, B' := \bigwedge X'$. Hence by B2, $B' \to C^\Box \wedge \Box B' \rhd^* [B']Z^\Box$. Note that we have $\Box B' \equiv \Box B^\Box$ and also $i\mathsf{K4} + \Box B' \vdash B' \leftrightarrow B^\Box$.

Now by using properties of $\rhd^*$ (A1-A3) and Lemma 3.14(2), we can deduce $(B^\Box \to C^\Box) \wedge \Box B^\Box \rhd^* [B^\Box]Z^\Box$. Then Lemma 3.14(1) implies $(B^\Box \to C^\Box) \wedge \Box B^\Box \rhd^* ([B]Z)^\Box$. Then by A1 and A2, we can deduce $((B \to C) \wedge \Box B)^\Box \rhd^* ([B]Z)^\Box$, as desired. ❑

The following theorem is analogous to the Theorem 4.18 in [2]:

Theorem 3.18. *For any modal proposition A, $A^\Box \rhd\lhd^* (A^\Box)^+$.*

Before proving this theorem, we state a corollary.

Corollary 3.19. $\mathsf{iH}^*_\sigma \vdash A^\Box \leftrightarrow (A^\Box)^-$.

Proof. Let $A^{\Box} = B(\Box C_1, \Box C_2, \ldots, \Box C_n)$ where $B(p_1, \ldots, p_n)$ is a non-modal proposition. It isn't hard to observe that for each $1 \leq j \leq n$, $i\mathsf{K4} \vdash \Box C_j \leftrightarrow \Box C_j^{\Box}$. By definition of $(A^{\Box})^-$, we have $(A^{\Box})^- = B(\Box C_1^+, \ldots, \Box C_n^+)$. Now by Lemma 3.7, we can deduce that $i\mathsf{K4} \vdash B(\Box C_1^+, \ldots, \Box C_n^+) \leftrightarrow B(\Box(C_1^{\Box})^+, \ldots, \Box(C_n^{\Box})^+)$. Then Theorem 3.18 implies that $\mathsf{iH}_\sigma^* \vdash \Box(C_i^{\Box})^+ \leftrightarrow \Box C_i^{\Box}$. Hence $\mathsf{iH}_\sigma^* \vdash (A^{\Box})^- \leftrightarrow A^{\Box}$. ❑

Proof. (**Theorem 3.18**) We prove by induction on $\mathfrak{o}(A^{\Box})$. Suppose that we have the desired result for each proposition B with $\mathfrak{o}(B^{\Box}) < \mathfrak{o}(A^{\Box})$. We treat A by the following cases.

1. (A1) A is atomic. Then $(A^{\Box})^+ = A^{\Box}$, by definition, and result holds trivially.

2. (A1-A4, B1) $A = \Box B, A = B \wedge C, A = B \vee C$. All these cases hold by induction hypothesis. In boxed case, we use of induction hypothesis and A4. In conjunction, we use of A1-A3 and in disjunction we use A1,A2 and B1.

3. $A = B \to C$. There are several sub-cases. similar to definition of NNIL-algorithm, an occurrence of a sub-formula B of A is said to be an *outer occurrence* in A, if it is neither in the scope of a $\Box$ nor in the scope of $\to$.

   (c).i.(A1-A3) C contains an outer occurrence of a conjunction. We can treat this case using induction hypothesis and TNNIL-algorithm.

   (c).ii.(A1-A3) B contains an outer occurrence of a disjunction. We can treat this case by induction hypothesis and TNNIL-algorithm.

   (c).iii. $B = \bigwedge X$ and $C = \bigvee Y$, where X, Y are sets of implications, atoms and boxed formulae. We have several sub-cases.

   (c).iii.α.(A1-A4, B3) X contains atomic variables. Let p be an atomic variable in X. Set $D := \bigwedge(X \setminus \{p\})$. Then

$$\begin{aligned}(A^{\Box})^+ &\equiv \boxdot[(p \wedge \Box p) \to (D^{\Box} \to C^{\Box})^+] \\ &\equiv \boxdot[(p \wedge \Box p) \to ((D \to C)^{\Box})^+]\end{aligned}$$

   On the other hand, we have by induction hypothesis and A1,A2 and B3, that

$$[(p \wedge \Box p) \to ((D \to C)^{\Box})^+] \bowtie^* (p \wedge \Box p) \to ((D \to C)^{\Box})$$

   which by use of A4 implies:

$$\Box[(p \wedge \Box p) \to ((D \to C)^{\Box})^+] \bowtie^* \Box[(p \wedge \Box p) \to ((D \to C)^{\Box})]$$

   And by use of A1-A3 we have

$$\boxdot[(p \wedge \Box p) \to ((D \to C)^{\Box})^+] \bowtie^* \boxdot[(p \wedge \Box p) \to ((D \to C)^{\Box})]$$

   Finally by A1 and A2 we have : $(A^{\Box})^+ \bowtie^* A^{\Box}$.

   (c).iii.β.(A1-A4, B3) X contains boxed formula. Similar to the previous case.

   (c).iii.γ.(A1, A2) X contains $\top$ or $\bot$. Trivial.

   (c).iii.δ.(A1-A4, B2, B3) X contains only implications. This case needs the axiom B2 and it seems to be the interesting case.

   By Lemma 3.15,

$$(A^{\Box})^+ \equiv \boxdot\left(\bigwedge_{E \to F \in X} \Box\left((E \to F)^{\Box}\right)^+ \to \left(\bigwedge\left\{\left((B \downarrow D \to C)^{\Box}\right)^+ \mid D \in X\right\} \wedge \left(([B]Z)^{\Box}\right)^+\right)\right).$$

Then by induction hypothesis, A1-A4 and B3 we have:

$$(A^{\Box})^{+} \bowtie^{*} \boxdot\left(\bigwedge_{E\to F\in X}\Box(E\to F)^{\Box}\to\left(\bigwedge\left\{(B\downarrow D\to C)^{\Box}\mid D\in X\right\}\wedge([B]Z)^{\Box}\right)\right)$$

$$\bowtie^{*}\left(\Box B\to\left(\bigwedge\{B\downarrow D\to C\mid D\in X\}\wedge[B]Z\right)\right)^{\Box}$$

We show that for each $E\in Z$,

$$(*)\qquad i\mathsf{K4}\vdash(\bigwedge\{(B\downarrow D)\to C\mid D\in X\}\wedge[B]E)\to A.$$

If $E=C$, we are done by $\mathsf{IPC}_{\Box}\vdash[B]C\to(B\to C)$. So suppose some $E\to F\in X$. We reason in $i\mathsf{K4}$. Assume $\bigwedge\{(B\downarrow D\to C\mid D\in X\}$, $[B]E$ and B. We want to derive C. We have $(\bigwedge(X\setminus\{E\to F\})\wedge F)\to C$, $[B]E$ and B. From B and $[B]E$, we derive E. Also from B, we derive $E\to F$, and so F. Hence we have $\bigwedge(X\setminus\{E\to F\})\wedge F$, which implies C, as desired.

Now (*) implies

$$i\mathsf{K4}\vdash\overbrace{(\bigwedge\{(B\downarrow D\to C\mid D\in X\}\wedge[B]Z)}^{G}\to A$$

Then by Proposition 3.13, we have $i\mathsf{K4}\vdash(G^{\Box}\wedge B^{\Box})\to C^{\Box}$. This implies $i\mathsf{K4}\vdash(B^{\Box}\to(G^{\Box}\wedge B^{\Box}))\to(B^{\Box}\to C^{\Box})$, and hence $i\mathsf{K4}\vdash(B^{\Box}\to G^{\Box})\to(B^{\Box}\to C^{\Box})$. Then because $B^{\Box}\to\Box B^{\Box}$, we have $i\mathsf{K4}\vdash(\Box(B^{\Box})\to G^{\Box})\to(B^{\Box}\to C^{\Box})$. Hence by necessitation, we derive $i\mathsf{K4}\vdash(\Box B\to(\bigwedge\{(B\downarrow D\to C\mid D\in X\}\wedge[B]Z))^{\Box}\to A^{\Box}$. Hence $(A^{\Box})^{+}\rhd^{*}A^{\Box}$.

To show the other way around, i.e., $A^{\Box}\rhd^{*}(A^{\Box})^{+}$, by Proposition 3.13, it is enough to show

$$A\rhd^{*}\left(\Box B\to\left(\bigwedge\{B\downarrow D\to C\mid D\in X\}\wedge[B]Z\right)\right)$$

or equivalently

$$A\wedge\Box B\rhd^{*}\left(\bigwedge\{B\downarrow D\to C\mid D\in X\}\wedge[B]Z\right)$$

We have $\mathsf{IPC}_{\Box}\vdash A\to\bigwedge\{B\downarrow D\to C\mid D\in X\}$, and hence by A1, $A\wedge\Box B\rhd^{*}\bigwedge\{B\downarrow D\to C\mid D\in X\}$. On the other hand, $A\wedge\Box B\rhd^{*}[B]Z$, which by A3, implies

$$A\wedge\Box B\rhd^{*}\left(\bigwedge\{B\downarrow D\to C\mid D\in X\}\wedge[B]Z\right)$$

❑

4 The Σ_1-Provability Logic of HA^*

In this section we will show that iH^*_{σ} is the provability logic of HA^* for Σ_1-substitutions.

Before we continue with the soundness and completeness theorem, let us state the main theorem from [2] that plays a crucial role in the sequel of this paper.

Theorem 4.1. *Let $A\in\mathsf{TNNIL}^{\Box}$ be a modal proposition such that $\mathsf{iGLC}\nvdash A$. Then there exists some arithmetical Σ_1-substitution σ such that $\mathsf{HA}\nvdash\sigma_{_{\mathsf{HA}}}(A)$.*

Proof. For the rather long proof of this fact, see [2], Theorems 4.26 and 5.1. ❑

4.1 The Soundness Theorem

Let us define some notions from [16]. We call a first-order sentence A, Σ-preserves B ($A \rhd_{T,\Sigma_1} B$), if for each Σ_1-sentence C, if $T \vdash C \to A$, then $T \vdash C \to B$. For modal propositions A and B, we define $A \rhd_{T,\Sigma_1,\Sigma_1} B$ iff for each arithmetical Σ-substitution σ_T, we have $\sigma_T(A) \rhd_{T,\Sigma_1} \sigma_T(B)$. For arbitrary modal sentences A, B, the notation $A \mathrel{|\!\!\sim}_{T,\Sigma_1} B$ means that $T \vdash \sigma_T(A)$ implies $T \vdash \sigma_T(B)$, for arbitrary Σ_1-substitution σ_T. All the above relations with a superscript of HA, means "an arithmetical formalization of that relation in HA ", for example, $A \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1} B$ means $\mathsf{HA} \vdash$ "$A \rhd_{\mathsf{HA}^*,\Sigma_1} B$".

Lemma 4.2. *1. For each first-order sentences A, B, $A \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1} B$ iff $A^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1} B^\Box$,*

2. For each propositional modal A, B, $A \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1} B$ iff $A^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} B^\Box$.

Proof. To prove part 1, use Lemma 2.6.1 and definitions of $\rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1}$ and $\rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1}$.

To prove part 2, note that $A \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1} B$ iff for all Σ-substitution σ, $\sigma_{\mathsf{HA}^*}(A) \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1} \sigma_{\mathsf{HA}^*}(B)$ iff for all Σ-substitution σ, $\sigma_{\mathsf{HA}^*}(A)^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1} \sigma_{\mathsf{HA}^*}(B)^\Box$ (by previous part) iff for all Σ-substitution σ, $\sigma_{\mathsf{HA}}(A^\Box) \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1} \sigma_{\mathsf{HA}}(B^\Box)$ iff $A^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} B^\Box$. ❏

Lemma 4.3. *$\rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1}$ is closed under $B1$.*

Proof. See [16], 9.1. ❏

Corollary 4.4. *$\rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1}$ is closed under $B1$.*

Proof. Immediate corollary of Lemma 4.2 and 4.3. ❏

Lemma 4.5. *$\rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1}$ satisfies $A1$-$A4$, $B1$, $B2'$ and $B3$.*

Proof. Proof of closure under A1-A4 and B3 is straightforward. Closure under B1 is by Lemma 4.3. For a proof of case B2$'$, see [16].9.2. ❏

Corollary 4.6. *$\rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1}$ satisfies $B2$.*

Proof. Let A, B, C, X, Z be as stated in defining B2. We must prove $A \wedge \Box B \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1} [B]Z$. Hence by Lemma 4.2, it is enough to show $(A \wedge \Box B)^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} ([B]Z)^\Box$. Let $X' := \{E^\Box \to F^\Box | E \to F \in X\}, B' := \bigwedge X', C' := C^\Box, Z' := \{E^\Box | E \to F \in X\} \cup \{C'\}$. Now Because $\rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1}$ satisfies B2$'$ (Lemma 4.5), we have $(B' \to C') \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} [B']Z'$. Note that $Z^\Box = Z'$ and $\mathsf{IPC}_\Box \vdash (B' \wedge \Box B') \leftrightarrow B^\Box$. Hence by Lemma 3.14.2, $i\mathsf{K4} + \Box B' \vdash [B']Z' \leftrightarrow [B^\Box]Z^\Box$. Also by Lemma 3.14.1, $i\mathsf{K4} + \Box B' \vdash [B^\Box]Z^\Box \leftrightarrow ([B]Z)^\Box$. So $i\mathsf{K4} + \Box B' \vdash [B']Z' \leftrightarrow ([B]Z)^\Box$. Now because $\rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1}$ satisfies A1, we have $\Box B' \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} [B']Z' \leftrightarrow ([B]Z)^\Box$. Now one can easily observe that because $\rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1}$ is closed under A1-A3, we can deduce $(B' \to C') \wedge \Box B' \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} ([B]Z)^\Box$. This by using A1-A3 implies $((B \to C) \wedge \Box B)^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} ([B]Z)^\Box$. Hence by Lemma 4.2.2, $(B \to C) \wedge \Box B \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} [B]Z$, as desired. ❏

Corollary 4.7. *$\rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1}$ is closed under $B3$.*

Proof. Let p be atomic or boxed and assume some A, B such that $A \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1} B$. Then by Lemma 4.2.2, $A^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} B^\Box$. Because $\rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1}$ satisfies B3, we get $p^\Box \to A^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} p^\Box \to B^\Box$. Now by A4, $\Box[p^\Box \to A^\Box] \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} \Box[p^\Box \to B^\Box]$, which implies $(p \to A)^\Box \rhd^{\mathsf{HA}}_{\mathsf{HA},\Sigma_1,\Sigma_1} (p \to B)^\Box$. Now by Lemma 4.2.2, $p \to A \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1} p \to B$, as desired. ❏

Lemma 4.8. *We have the following inclusions:*

$$\blacktriangleright^* \subseteq \rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1} \subseteq \mathrel{|\!\!\sim}^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1}$$

Proof. The second inclusion is a trivial. We only prove the first inclusion. We show that $\rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1}$ is closed under A1-A4 and B1-B3. One can observe that $\rhd^{\mathsf{HA}}_{\mathsf{HA}^*,\Sigma_1,\Sigma_1}$ is closed under A1-A4 and we leave this to the reader. Closure under B1, B2 and B3 is by Corollaries 4.4,4.6 and 4.7, respectively. ❑

Theorem 4.9. (**Soundness**) *iH^*_σ is sound for Σ_1-arithmetical interpretations in HA^*, i.e. $\mathsf{iH}^*_\sigma \subseteq \mathcal{PL}_\sigma(\mathsf{HA}^*)$.*

Proof. We show that for arbitrary Σ-substitution, σ_{HA^*}, and for any A, if $\mathsf{iH}^*_\sigma \vdash A$, then $\mathsf{HA}^* \vdash \sigma_{\mathsf{HA}^*}(A)$. This can be done by induction on the complexity of $\mathsf{iH}^*_\sigma \vdash A$. All inductive steps clearly holds, except for the axioms $\Box A \to \Box B$ with $A \blacktriangleright^* B$. This case is a direct consequence of Lemma 4.8. ❑

4.2 The Completeness Theorem

Theorem 4.10. *Σ_1-arithmetical interpretations in HA^* are complete for iH^*_σ, i.e.*

$$\mathcal{PL}_\sigma(\mathsf{HA}^*) \subseteq \mathsf{iH}^*_\sigma$$

Proof. We prove the Completeness Theorem contra-positively. Let $\mathsf{iH}^*_\sigma \nvdash A(p_1, \ldots, p_n)$. Then $\mathsf{iH}^*_\sigma \nvdash A^\Box$ and hence by Corollary 3.19, $\mathsf{iH}^*_\sigma \nvdash (A^\Box)^-$. This, by Theorem 3.6, implies $\mathsf{iH}^*_\sigma \nvdash ((A^\Box)^-)^*$ and hence $\mathsf{iH}^*_\sigma \nvdash (A^\Box)^+$, and a fortiori, $\mathsf{iGLC} \nvdash (A^\Box)^+$. Hence by Theorem 4.1, there exists some Σ_1-substitution σ, such that $\mathsf{HA} \nvdash \sigma_{\mathsf{HA}}((A^\Box)^+)$. Hence by Lemma 3.7.1, $\mathsf{HA} \nvdash \sigma_{\mathsf{HA}}(A^\Box)$ and by Lemma 2.7, $\mathsf{HA}^* \nvdash \sigma_{\mathsf{HA}^*}(A)$. ❑

Corollary 4.11. *For any modal proposition A, $\mathsf{iH}^*_\sigma \vdash A$ iff $\mathsf{iH}_\sigma \vdash \mathsf{A}^\Box$.*

Proof. By Theorems 4.9 and 4.10 and Lemma 2.7. ❑

Corollary 4.12. *iH^*_σ is decidable.*

Proof. A proof can be given either with inspections in the proof of the Completeness Theorem (4.10) or by using the decidability of iH_σ [2] and Corollary 4.11. ❑

Open problems

1. The statement of Corollary 4.11 is *purely propositional*. However, our proof of this corollary is based on Theorem 4.10, that has *arithmetical* theme. A tempting problem is to find a *direct propositional proof* for this corollary. Then we can derive Theorem 4.10.

2. We conjecture that the full provability logic of HA^* is the logic iH^*, axiomatized as follows

$$\mathsf{iH}^* := \mathsf{iGL} + \mathsf{CP} + \{\Box A \to \Box B : A \blacktriangleright^*_\alpha B\},$$

   in which the relation $\blacktriangleright^*_\alpha$ is defined as the smallest relation satisfying:

   - A1. If $i\mathsf{K4} \vdash A \to B$, then $A \blacktriangleright^*_\alpha B$,
   - A2. If $A \blacktriangleright^*_\alpha B$ and $B \blacktriangleright^*_\alpha C$, then $A \blacktriangleright^*_\alpha C$,
   - A3. If $C \blacktriangleright^*_\alpha A$ and $C \blacktriangleright^*_\alpha B$, then $C \blacktriangleright^*_\alpha A \wedge B$,
   - A4. If $A \blacktriangleright^*_\alpha B$, then $\Box A \blacktriangleright^*_\alpha \Box B$,
   - B1. If $A \blacktriangleright^*_\alpha C$ and $B \blacktriangleright^*_\alpha C$, then $A \vee B \blacktriangleright^*_\alpha C$,
   - B2. Let X be a set of implications, $B := \bigwedge X$ and $A := B \to C$. Also assume that $Z := \{E | E \to F \in X\} \cup \{C\}$. Then $A \wedge \Box B \blacktriangleright^*_\alpha \{B\}Z$,
   - B3. If $A \blacktriangleright^*_\alpha B$, then $\Box C \to A \blacktriangleright^*_\alpha \Box C \to B$.

The notation $\{A\}(B)$, for modal propositions A and B, is defined inductively:

- $\{A\}(\Box B) = \Box B$ and $\{A\}(\bot) = \bot$.
- $\{A\}(B_1 \circ B_2) = \{A\}(B_1) \circ \{A\}(B_2)$, for $\circ \in \{\vee, \wedge\}$,
- $\{A\}(B) = A \to B$ for all of the other cases, i.e. when B is atomic variable or implication.

And $\{A\}\Gamma$, for a set Γ of modal propositions, is defined as $\bigvee_{B\in\Gamma}\{A\}(B)$.

References

[1] M. Ardeshir and M. Mojtahedi, *Reduction of provability logics to Σ_1-provability logics*, Logic Journal of IGPL **23** (2015), no. 5, 842–847.

[2] ______, *The Σ_1-Provability Logic of* HA, Annals of Pure and Applied Logic **169** (2018), no. 10, 997–1043.

[3] S. Artemov and L. Beklemishev, *Provability logic*, in Handbook of Philosophical Logic (D. Gabbay and F. Guenthner, eds.), vol. 13, Springer, 2nd ed., 2004, pp. 189–360.

[4] H. Friedman, *The disjunction property implies the numerical existence property*, Proc. Nat. Acad. Sci. U.S.A. **72** (1975), no. 8, 2877–2878. MR 0379141 (52 #47)

[5] K. Gödel, *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I*, Monatsh. Math. Phys. **38** (1931), no. 1, 173–198.

[6] ______, *Eine interpretation des intuitionistischen aussagenkalkuls*, Ergebnisse eines mathematischen Kolloquiums **4** (1933), 39–40, English translation in: S. Feferman etal., editors, Kurt Gödel Collected Works, Vol. 1, pages 301-303. Oxford University Press, 1995.

[7] R. Iemhoff, *Provability logic and admissible rules*, Ph.D. thesis, University of Amsterdam, 2001.

[8] D. Leivant, *Absoluteness in intuitionistic logic*, Ph.D. thesis, University of Amsterdam, 1975.

[9] ______, *Absoluteness of intuitionistic logic*, Mathematical Centre Tracts, vol. 73, Mathematisch Centrum, Amsterdam, 1979.

[10] M. Löb, *Solution of a Problem of Leon Henkin*, Journal of Symbolic Logic **20** (1955), no. 2, 115–118.

[11] J. Myhill, *A note on indicator-functions*, Proceedings of the AMS **39** (1973), 181–183.

[12] R. M. Solovay, *Provability interpretations of modal logic*, Israel J. Math. **25** (1976), no. 3-4, 287–304.

[13] A. S. Troelstra and D. van Dalen, *Constructivism in mathematics. Vol. I*, Studies in Logic and the Foundations of Mathematics, vol. 121, North-Holland Publishing Co., Amsterdam, 1988, An introduction.

[14] A. Visser, *Aspects of diagonalization and provability*, Ph.D. thesis, Utrecht University, 1981.

[15] ______, *On the completeness principle: a study of provability in Heyting's arithmetic and extensions*, Ann. Math. Logic **22** (1982), no. 3, 263–295.

[16] ______, *Substitutions of Σ_1^0 sentences: explorations between intuitionistic propositional logic and intuitionistic arithmetic*, Ann. Pure Appl. Logic **114** (2002), no. 1-3, 227–271, Commemorative Symposium Dedicated to Anne S. Troelstra (Noordwijkerhout, 1999).

[17] A. Visser, J. van Benthem, D. de Jongh, and G. R. R. de Lavalette, NNIL*, a study in intuitionistic propositional logic*, Modal logic and process algebra (Amsterdam, 1994), CSLI Lecture Notes, vol. 53, CSLI Publ., Stanford, CA, 1995, pp. 289–326.

[18] A. Visser and J. Zoethout, *Provability Logic and the Completeness Principle*, ArXiv e-prints (2018).