Logic of Proofs[*]

Sergei Artëmov

Steklov Mathematical Institute,
Vavilov str. 42,
117966 Moscow, Russia
email: art@log.mian.su[†]

August 10, 1993

Abstract

In this paper individual proofs are integrated into provability logic. Systems of axioms for a logic with operators "A is provable" and "p is a proof of A" are introduced, provided with Kripke semantics and decision procedure. Completeness theorems with respect to the arithmetical interpretation are proved.

1 Introduction

In [1] and [2] proofs were incorporated into propositional logic by means of labeled modalities. The basic labeled modal logic contains the propositional logic enriched by unary operators $\Box_{p_i}$, $i = 0, 1, 2, \ldots$. This language helps to provide a logical treatment of a rather general situation when we are interested not only to know that a certain statement A is valid, but also have to keep track on some evidences of its validness: $\Box_p A$ may stand for "p is a proof of A", "p is a program which computes A", "A has a proof of the complexity p", etc. The language of the provability logic ([3]) with the *provability* operator $\Box$ only, where $\Box A$ stands for "A is provable", can not do this job. However, labeled modalities alone fail to express some key principles of provability, e.g. the fact that a set of theorems is closed under *modus ponens*; it can easily be done with the use of the provability operator by the axiom scheme $\Box(A \to B) \to (\Box A \to \ \Box B)$. In the

[*] *Annals of Pure and Applied Logic*, v. 67, pp. 29-59, 1994.

[†]Supported by the grant # 93-011-16015 of the Russian Foundation for Fundamental Research, and by the US Army Research Office through the Mathematical Sciences Institute of Cornell University, Contract DAAL03-91-C-0027

current paper (as well as in [4] and [5]) we consider a propositional language, which together with labeled modalities for proofs contains also the "usual" modal operator $\Box$ for provability. In the context of proof interpretations it is probably the minimal adequate language. Here is an example of how the proof operators extend the expressive power of provability logic: because of the Godel Incompleteness Theorems, provability logic cannot either prove or reject the formula $\neg\Box\bot$, which is a modal equivalent of the consistency statement. But the logic of proofs derives formulas $\neg\Box_p\bot$, $\neg\Box_p\Box\bot$, which are natural labeled modal formulations of the unprovability of $\bot$.

The Kripke completeness proofs demonstrate that the basic axioms of the labeled modal logic: $\Box_p A \to A$, $\Box_p A \to \Box\Box_p A$ and $\neg\Box_p A \to \Box(\neg\Box_p A)$ are compatible with the Kripke semantics. Moreover, the extension of a usual modal logic by these new axioms does not destroy the corresponding Kripke frames.

A labeled modal language $\mathcal{L}^+$ contains two sorts of variables, $p_0, p_1, \ldots$ (called *proof variables*) and $S_0, S_1, \ldots$ (called *sentence variables*), symbol $\to$ for the classical implication, the truth value $\bot$ for absurdity (the usual Boolean connectives, and the truth value $\top$ for truth are defined as abbreviations), the usual modality $\Box$, and for each proof variable p_i the unary modal operator $\Box_{p_i}(\cdot)$ (labeled modality). The set of formulas of $\mathcal{L}^+$ is thus generated from the **atomic** formulas $\bot, S_0, S_1, \ldots$ by $\to$ as usual, and by the modal operators as follows: if A is an $\mathcal{L}^+$ formula, then $\Box A$ is an $\mathcal{L}^+$ formula, if p is a proof variable and A an $\mathcal{L}^+$ formula, then $\Box_p A$ is an $\mathcal{L}^+$ formula; we call these formulas **quasiatomic**, or **q-atomic** for short. In the sequel under a modal formula we understand a formula in the language $\mathcal{L}^+$. We use small letters $p, q, r, \ldots$ for proof variables, capital letters $S, T, \ldots$ for sentence variables and $A, B, C, \ldots$ for modal formulas. Let $\mathcal{L}$ denote the usual modal language over $\bot, S_0, S_1, \ldots$ with the only modality $\Box$, i.e. $\mathcal{L}$ is a labeled-modalities-free fragment of $\mathcal{L}^+$.

Here we make some adaptation of the common unification technique (cf.[6]) to the language $\mathcal{L}^+$. Those readers who are familiar with substitutions, most general unifiers, etc., may go directly to the definition 1.2.

A **substitution** θ is a finite set of the form

$$(T_1 \leftarrow A_1, \ldots T_n \leftarrow A_n, q_1 \leftarrow r_1, \ldots q_m \leftarrow r_m),$$

where the T_i are distinct sentence variables, the q_j are distinct proof variables and each A_i is a modal formula other than T_i, each r_j is a proof variable other than q_j. For convenience we say an **expression** is any modal formula or a proof variable. For E an expression, we write $E\theta$ for the result of simultaneously replacing each occurrence of T_i in E by A_i and each occurrence of q_j in E by r_j for every $i \leq n$, $j \leq m$. If X is a set of expressions, then

$$X\theta \;=\; \bigcup_{E\in X} E\theta.$$

If τ and θ are substitutions, then $\tau = \theta$ iff $F\tau \equiv F\theta$[1] for every modal formula F. If $\theta = (T_1 \leftarrow A_1, \ldots, q_1 \leftarrow r_1, \ldots)$ and $\sigma = (T'_1 \leftarrow A'_1, \ldots, q'_1 \leftarrow r'_1, \ldots)$, then a **composition** $\theta \circ \sigma$ (or just $\theta\sigma$ for short) of θ and σ is

$$(T_1 \leftarrow A_1\sigma, \ldots, q_1 \leftarrow r_1\sigma, \ldots, T'_1 \leftarrow A'_1, \ldots, q'_1 \leftarrow r'_1, \ldots),$$

less any $T_i \leftarrow A_i\sigma$ for which $T_i \equiv A_i\sigma$, and any $q_j \leftarrow r_j\sigma$ for which $q_j \equiv r_j\sigma$, any $T'_i \leftarrow A'_i, q_j\prime \leftarrow r'_j$ for which $T'_i, q'_j \in \{T_1, \ldots, q_1, \ldots\}$. Composition is naturally defined to satisfy $(E\theta)\sigma = E(\theta\sigma)$ and $(\psi\theta)\sigma = \psi(\theta\sigma)$ for any expression E and substitutions ψ, θ and σ.

If $X = \{E_1, \ldots E_n\}$ is a set of expressions we say a substitution θ is a **unifier for** X if $E_1\theta \equiv E_2\theta \equiv \ldots E_n\theta$, i.e. $X\theta$ is a singleton. X is said to be unifiable if it has a unifier. Below, speaking about the unification of $B_1, \ldots, B_n$, we'll mean the unification of the set $\{B_1, \ldots, B_n\}$. With formulas $A_1 \to B_1, A_2 \to B_2$ to be unified we associate a **set of equations** $\{A_1 = A_2,\ B_1 = B_2\}$, with formulas $\Box A, \Box B$ a set of equations $\{A = B\}$, with formulas $\{\Box_p A, \Box_q B\}$ a set of equations $\{A = B,\ p = q\}$. A substitution θ such that $E_1\theta \equiv F_1\theta, \ldots, E_n\theta \equiv F_n\theta$ is called a **unifier of the set of equations** $\{E_1 = F_1, \ldots, E_n = F_n\}$. Sets of equations are called **equivalent** if they have the same unifiers. A (possibly empty) set of equations is called **solved** if it is of the form

$$T_1 = A_1, \ldots T_n = A_n, q_1 = r_1, \ldots q_m = r_m,$$

where T_i and q_j are distinct variables and none of them occurs in a right hand side of any equation. Such solved set of equations determines the substitution

$$\theta \ = \ (T_1 \leftarrow A_1, \ldots T_n \leftarrow A_n, q_1 \leftarrow r_1, \ldots q_m \leftarrow r_m).$$

This substitution θ is a unifier of this set of equations and clearly it is its **most general unifier (mgu)**, i.e. for any other unifier σ of this set of equations there exists a substitution λ such that $\sigma = \theta\lambda$. Thus to find an mgu of a set X of expressions it suffices to transform the associated set of equations into an equivalent one which is solved.

Unification Algorithm. Nondeterministically choose from the set of equation an equation of a form below and perform the associated action.

1. $A_1 \to A_2 = B_1 \to B_2$: replace by the equations $A_1 = B_1,\ A_2 = B_2$,
   $\Box A = \Box B$: replace by the equation $A = B$,
   $\Box_p A = \Box_q B$: replace by the equations $A = B,\ p = q$;
2. equation of formulas of sort $A \to B,\ \Box A,\ \Box_p A,\ \bot$ with *different* principal connectives: halt with failure;
3. $\bot = \bot$, $S_i = S_i$, or $p_i = p_i$: delete the equation;

[1] Here $\equiv$ denotes a syntactical identity.

4. $A = S_i$ where A is not a variable: replace by the equation $S_i = A$;

5. $S_i = A$ where $S_i \not\equiv A$ and S_i has another occurrence in the set of equations: if S_i appears in A, then halt with failure, otherwise replace S_i by A in every other equation;

6. $p_i = p_j$ where $i \neq j$ and p_i has another occurrence in the set of equations: replace p_i by p_j in every other equation.

Algorithm terminates when no step can be performed or when failure arises.

The Unification Algorithm for a set X succeeds iff X is unifiabe, and it produces a special most general unifier θ such that

1.1 Lemma. (Cf.[6]) $Dom(\theta) \cap Val(\theta) = \emptyset$ and θ is idempotent, i.e. $\theta\theta = \theta$.

1.2 Definition. For a convenience we consider some deterministic variant $\mathbf{U}$ of the Unification Algorithm by fixing an order of the equations for this algorithm to choose. Let τ_{AB} be the mgu of A, B obtained by $\mathbf{U}$ on the set of equations associated with the pair of formulas $\{A, B\}$. We call τ_{AB} the **standard** mgu of A, B.

1.3 Definition. Let $\quad C = D \pmod{A = B} \quad$ be an abbreviation for

$$\text{“for every substitution } \theta \;\; (A\theta \equiv B\theta \implies C\theta \equiv D\theta)\text{”}.$$

Apparently, if A, B are not unifiable, then $C = D \pmod{A = B}$ holds for all C and B.

1.4 Lemma. If A and B are unifiable, then

$$C = D \pmod{A = B} \quad \iff \quad C\tau_{AB} \equiv D\tau_{AB}.$$

Proof. Direction ($\Longrightarrow$) is obvious as τ_{AB} unifies A and B. Direction ($\Longleftarrow$): let $C\tau_{AB} \equiv D\tau_{AB}$ and θ be an arbitrary unifier A and B. As τ_{AB} is an mgu of A, B for some λ we have $\theta = \tau_{AB} \circ \lambda$, then

$$C\theta \equiv C\tau_{AB} \circ \lambda \equiv D\tau_{AB} \circ \lambda \equiv D\theta.$$

■

1.5 Corollary. The relation $\quad C = D \pmod{A = B} \quad$ is decidable.

In what follows we assume, for short, the Peano Arithmetic $\mathtt{PA}$ to be the basic theory for proof and provability predicates. We do not restrict ourselves by considering *a priori* a specific proof predicate; the proof logic of such a predicate may depend on occasional details of numerations. However, the logic of the usual Gödel proof predicate is axiomatized in Chapter 4. We denote the usual Gödel proof predicate as $\mathit{Proof}(x,y)$ and the usual provability predicate as $\mathit{Provable}(y)$, i.e. $\mathit{Provable}(y)$ coincides with $\exists x \mathit{Proof}(x,y)$.

1.6 Definition. An arithmetical formula $\mathit{Prf}(x,y)$ is called a **standard proof predicate** (cf.[7]) iff

1. $\mathit{Prf}(x,y)$ is equivalent in $\mathtt{PA}$ to a recursive formula;
2. $\mathit{Prf}(x,y)$ numerates the theorems of $\mathtt{PA}$:

$$\mathtt{PA} \vdash \varphi \iff \text{for some } n \in \omega \ \mathit{Prf}(n, \ulcorner\varphi\urcorner) \text{ is true}^2,$$

3. the formula $\mathit{Pr}(y) := \exists x \mathit{Prf}(x,y)$ satisfies:

$$\mathtt{PA} \vdash (\mathit{Pr}(\ulcorner A\urcorner) \& \mathit{Pr}(\ulcorner A \rightarrow B\urcorner)) \rightarrow \mathit{Pr}(\ulcorner B\urcorner),$$

$$\mathtt{PA} \vdash \sigma \rightarrow \mathit{Pr}(\ulcorner\sigma\urcorner)$$

for every Σ_1^0 arithmetical sentence σ.

We choose a fairly general definition of a standard proof predicate; it allows not only proofs as first argument, but also programs, or other special codes. On the other hand all specific **provability** predicates $\mathit{Pr}(y) := \exists x\ \mathit{Prf}(x,y)$, that we deal with in this paper are (provably in $\mathtt{PA}$) equivalent to the usual one $\mathit{Provable}(y)$. Of course, if $\mathit{Prf}(x,y)$ is a recursive formula and

$$\mathtt{PA} \vdash \forall y(\mathit{Pr}(y) \ \leftrightarrow \ \mathit{Provable}(y)),$$

then $\mathit{Prf}(x,y)$ is a standard proof predicate.

1.7 Definition. Let Prf be a standard proof predicate and let ϕ be a function which assigns to each proof variable p some $n \in \omega$ and to each sentence variable S a sentence of $\mathtt{PA}$. An **arithmetical interpretation** $*$ is a pair (Prf, ϕ). The arithmetical translation A^* of a modal formula A under the interpretation $*$ is the extension of ϕ to all modal formulas by: $\bot^* := (0=1)$, $p^* := \phi(p)$ for a proof variable p, $S^* := \phi(S)$ for a sentence variable S, $(\cdot)^*$ commutes with the Boolean connectives,

$$(\Box A)^* := \mathit{Pr}(\ulcorner A^* \urcorner),$$

[2] In this paper we do not distinguish between the natural number n and its numeral $\overline{n}$. As usual, $\ulcorner\varphi\urcorner$ denotes the Gödel number of φ.

and

$$(\Box_p A)^* := \mathit{Prf}\,(p^*, \ulcorner A^{*}\urcorner).$$

If $* = (\mathit{Prf}, \varphi)$ is an arithmetical interpretation of a labeled modal language $\mathcal{L}$ and θ a substitution, then their composition $\sharp = \theta \circ *$ (or just $\theta *$ for short) is an arithmetical interpretation of $\mathcal{L}$ with the same proof predicate such that $v^\sharp = (v\theta)^*$ for v a sentence or a proof variable. The equality of two interpretations $* = \sharp$ is obviously equivalent to

$$\text{“for every modal formula } F \quad (F^* \equiv F^\sharp)\text{”}.$$

1.1 Axioms of the Logic of Proofs

The following system $\mathcal{B}$ will provide a complete axiomatization of the class of all standard proof predicates (Theorem 2.12).

1.8 Definition. Axioms of $\mathcal{B}$:

(A1)	Boolean tautologies in the language $\mathcal{L}^+$	
(A2)	$\Box(A \to B) \to (\Box A \to \Box B)$	distributivity
(A3)	$\Box(\Box A \to A) \to \Box A$	Löb axiom
(A4)	$\Box_p A \to A$	q-reflexivity
(A5)	$\Box_p A \to \Box\Box_p A$	stability
(A6)	$\neg\Box_p A \to \Box(\neg\Box_p A)$	stability

where p is a proof variable, A and B are formulas.

Rules of $\mathcal{B}$:

(R1) $$\frac{A \qquad A \to B}{B}$$

(R2) $$\frac{A}{\Box A}$$

(R3) $$\frac{\Box A}{A}.$$

1.9 Definition. A standard proof predicate Prf is **functional** if for all $l, m, n \in \omega$

$$\mathit{Prf}\,(l,m),\ \mathit{Prf}\,(l,n) \quad \Longrightarrow \quad m = n.$$

An arithmetical interpretation (Prf, ϕ) is called functional iff Prf is a functional proof predicate. The system $\mathcal{F}$ below will axiomatize the class of all functional proof predicates (Theorem 3.14).

1.10 Definition. The system $\mathcal{F}$ is $\mathcal{B}$ together with the **functionality axiom**:

$$\textbf{(A7)} \quad \Box_p A \& \Box_p B \to (C \to D) \ \text{ if } \ C = D \quad (\text{mod } A = B),$$

where p is a proof variable, A, B, C, D are formulas.

1.11 Comment. If A, B are not unifiable, then $\mathcal{F} \vdash \neg(\Box_p A \& \Box_p B)$. Indeed, put $C := \top$ and $D := \bot$ in **A7**. The following scheme called the **local functionality principle** is provable in $\mathcal{F}$:

$$\Box_p A \& \Box_p B \to (C \leftrightarrow C\tau_{AB}).$$

Indeed, as τ_{AB} is idempotent $C\tau_{AB}\tau_{AB} \equiv C\tau_{AB}$ and $C = C\tau_{AB} \pmod{A = B}$; thus this principle is a special case of the functionality axiom. Moreover, one can even replace **A7** in the list of axioms of $\mathcal{F}$ by this local functionality principle: indeed, by

$$\Box_p A \& \Box_p B \to (C \leftrightarrow C\tau_{AB})$$

and

$$\Box_p A \& \Box_p B \to (D \leftrightarrow D\tau_{AB})$$

one can easily infer

$$\Box_p A \& \Box_p B \to (C \to D).$$

(One has only to take care of the case when A, B are not unifiable). In fact the first formulation of $\mathcal{F}$ came with the local functionality as an axiom. The author is grateful to Tyko Strassen who suggested a substitution-free form of the functionality axiom and argued strongly for it.

1.12 Definition. The system $\mathcal{M}$ (after *Monotone*) will axiomatize the usual Gödel proof predicate (Theorem 4.7). $\mathcal{M}$ is $\mathcal{F}$ together with the **monotonicity axiom**:

$$\textbf{(A8)} \quad \neg[\Box_{p_1} A_2(p_2) \& \Box_{p_2} A_3(p_3) \& \ldots \& \Box_{p_n} A_1(p_1)],$$

where $A_i(p_i)$, $(i = 1, \ldots n)$ is a formula in which the proof variable p_i occurs.

The main results of the current paper one can find in technical reports [4] and [5].

2 Logic of standard proof predicates

The language $\mathcal{L}^+$ is able to distinguish some standard proof predicates; for example the formula $\neg\Box_p\Box_p\top$, which claims than p cannot be simultaneously a proof of $\top$ and a proof of $\Box_p\top$, is true for a functional Prf, but false for some other standard proof predicates (see below). Thus, there is no one logic which would be complete with respect to each single standard proof predicate. In this section we prove that $\mathcal{B}$ is complete with respect to the entire class of standard proof predicates.

2.1 System $\mathcal{B}^-$

2.1 Definition. For some technical reasons we consider an auxiliary system $\mathcal{B}^-$ which has the same axioms as $\mathcal{B}$ and rules **R1** and **R2**.

Axioms **A1-A3** came from the Gödel–Löb provability logic $\mathtt{GL}$ (cf.[3],[8],[9]) together with rules **R1** and **R2**. So $\mathcal{B}^-$ derives all $\mathtt{GL}$-tautologies. Axiom **A4** came from the system $\mathcal{P}$ ([1]), **A5** and **A6** reflect how $\Box$ and $\Box_{p_i}(\cdot)$ match together.

2.2 Example. $\mathcal{B}^- \vdash \Box_p A \to \Box A$ for every $A \in \mathcal{L}^+$. Indeed,

1. $\Box_p A \to A$ (Axiom **A4**)
2. $\Box\Box_p A \to \Box A$ (From 1. by **A1, A2** and **R1, R2**)
3. $\Box_p A \to \Box\Box_p A$ (Axiom **A6**)
4. $\Box_p A \to \Box A$ (From 2. and 3.)

The difference between $\mathcal{B}^-$ and $\mathcal{B}$ can be demonstrated by the following example.

2.3 Example. $\mathcal{B}$ proves $\neg\Box_p\Box\bot$ but $\mathcal{B}^-$ does not.

Here is a derivation of $A := \neg\Box_p\Box\bot$ in $\mathcal{B}$:

1. $\bot \to A$ (tautology **A1**)
2. $\Box\bot \to \Box A$ (from 1. by **A1, A2** and **R1, R2**)
3. $(\neg\Box\bot \to \Box A) \to \Box A$ (from 2. by **A1** and **R1**)
4. $[(\neg\Box\bot \to A) \wedge (A \to \Box A)] \to \Box A$ (from 3. by **A1** and **R1**)
5. $\neg\Box\bot \to A$ (is equivalent to an **A4** axiom $\Box_p\Box\bot \to \Box\bot$)
6. $A \to \Box A$ (axiom **A6**)

7. $\Box A$ (from 4,5 and 6)
8. A (rule **R3**)

We will show that $\mathcal{B}^- \nvdash \neg\Box_p\Box\bot$ below in Corollary 2.7 after introducing semantics for $\mathcal{B}^-$.

2.2 Kripke semantics for $\mathcal{B}^-$ and $\mathcal{B}$

There are two possible ways to build a Kripke semantics here: either to reduce the problem to that for the "host" modal logic $\mathtt{GL}$ by means of the adequate sets trick, or to proceed the standard canonical model proof taking care of new features generated by the labeled modalities. Either of these ways succeeds, and either has its advantages. In order to learn more about labeled modal logics we choose the first of these approaches for $\mathcal{B}$ and the second one for $\mathcal{F}$ and $\mathcal{M}$.

2.4 Definition. A **frame** is a pair $(K, \prec)$, where $K \neq \emptyset$ and K is finite, $\prec$ is an irreflexive tree-like ordering of K[3]. In the sequel we call elements of K *nodes*, and let $\succ$ stand for $(\prec)^{-1}$. A **model** is a triple $(K, \prec, \Vdash)$, where $(K, \prec)$ is a frame and $\Vdash$ is a forcing relation between nodes and $\mathcal{L}^+$ formulas satisfying the following **forcing conditions**:

1. $\Vdash$ respects Boolean operations at each node,
2. $x \Vdash \Box\varphi$ iff $\forall y \succ x \; y \Vdash \varphi$,
3. for every q-atomic formula $\Box_p A$
   $\forall x \in K \; x \Vdash \Box_p A$ or $\forall x \in K \; x \Vdash \neg\Box_p A$,
   (stability)
4. for every q-atomic formula $\Box_p A$
   $x \Vdash \Box_p A \implies x \Vdash A$,
   (q-reflexivity)

Note that it may be the case when $A \leftrightarrow B$ is a Boolean tautology (A and B are modal formulas), but $x \Vdash \Box_p A$ and $x \Vdash \neg\Box_p B$. We say that a modal formula A is **valid in a model** $\mathcal{K} = (K, \prec, \Vdash)$ if A holds at every node; and $\mathcal{K}$ is a **countermodel to** A if A is not valid in $\mathcal{K}$, i.e. $\neg A$ holds at some node $x \in K$. For any $A \in \mathcal{L}^+$, let $\mathbf{H}(A) = \bigwedge\{\Box B \to B \mid \Box B \text{ is a subformula of } A\}$. We call a model A**-sound** if its root satisfies $\mathbf{H}(A)$.

2.5 Theorem. $\mathcal{B} \vdash A \iff A$ holds in all A-sound models.

[3] In fact any $\mathtt{GL}$-frame (i.e. transitive reverse well-founded) would fit here, but we choose finite trees going directly to a strong form of the Kripke completeness theorem for $\mathcal{B}$ which will be useful for the arithmetical completeness proof.

Proof. We will first prove the Kripke completeness of $\mathcal{B}^-$.

2.6 Lemma. For every modal formula A

$$\mathcal{B}^- \vdash A \quad \Longrightarrow \quad A \text{ is valid in all models.}$$

Proof. Correctness ($\Longleftarrow$) follows easily from an observation that $(K, \prec)$ is a $\mathtt{GL}$ frame; the correctness of **A4-A6** is immediate by the stability and q-reflexivity forcing conditions.

2.7 Corollary. $\neg\Box_p\Box\bot$ is not derivable in $\mathcal{B}^-$.

Indeed, consider a model $\mathcal{K} = (K, \prec, \Vdash)$, where $K = \{a\}$ is a singleton, $\prec$ is thus empty, $a \Vdash \Box_p\Box\bot$, and none of other atomic and q-atomic formulas is true in $\mathcal{K}$. Clearly, $\mathcal{K}$ is a model, because $a \Vdash \Box\bot$ and thus the q-reflexivity property is maintained.

We proceed now with the side "$\Longleftarrow$" of Lemma 2.6.

2.8 Definition. Let SbA stand for the set of all subformulas of A. A set X of modal formulas is **adequate** if it is closed under subformulas. If X is an adequate set, then a triple $\mathcal{K} = (K, \prec, \Vdash)$ is an X**-model** if the forcing relation $\Vdash$ is defined only for formulas from X and for these formulas all the forcing conditions 1-4 (definition 2.4) are respected.

Note that every X-model can be extended to a model by defining $a \nVdash \varphi$ for each node $a \in K$ and each atomic and q-atomic formula $\varphi \notin X$. Under this procedure the stability and q-reflexivity forcing conditions are clearly preserved.

We prove now that for every modal formula A and every finite adequate set X containing A

$$\mathcal{B}^- \nvdash A \quad \Longrightarrow \quad \text{there is an } X\text{-countermodel for } A.$$

Suppose A is a modal formula and $\mathcal{B}^- \nvdash A$. Let X be a finite adequate set containing A,

$$D_0 = \Box_{q_0}A_0, \ldots, D_n = \Box_{q_n}A_n$$

be the list of all q-atomic formulas of X, and let $T_0, \ldots, T_n$ be sentence variables not occurring in X. To every $B \in X$ we associate an $\mathcal{L}$-formula B^t such that B is the result of substituting all those occurrences of D_i which are not in scopes of any labeled modalities for T_i throughout B^t for all i with $0 \leq i \leq n$. Thus the following holds:

- B^t is a formula in the language $\mathcal{L}$,
- B is the result of substituting T_i for D_i for all i with $0 \leq i \leq n$.

Let Y be a set of $\mathcal{L}$ formulas, consisting of

1. $T_i \to \Box T_i$,
2. $\neg T_i \to \Box\neg T_i$,
3. $T_i \to A_i^t$,
4. $\Box(T_i \to A_i^t)$

for $0 \le i \le n$.

It is clear, that the substitution T_i for D_i makes Y some set of theorems of $\mathcal{B}^-$, and since $\mathcal{B}^- \nvdash A$ it follows that $\mathtt{GL} \nvdash \bigwedge Y \to A^t$. The Kripke style completeness theorem for $\mathtt{GL}$ guarantees that there is a $\mathtt{GL}$-model $\mathcal{K} = (K, \prec, \Vdash)$, satisfying $\bigwedge Y \wedge \neg A^t$ at its root. Now keep the frame $(K, \prec)$ and change $\Vdash$ to $\Vdash'$, putting

$$a \Vdash' G \quad\Longleftrightarrow\quad a \Vdash G^t,$$

for $G \in X$. The resulting triple $\mathcal{K}' = (K, \prec, \Vdash')$ is an X-model, whose root satisfies $\neg A$. Indeed, we have only to check the stability and q-reflexivity forcing conditions on $\mathcal{K}'$ for q-atomic formulas from X.

Let a be the root node of the frame $(K, \prec)$; if $a \Vdash T_i$, $\ (0 \le i \le n)$, then $a \Vdash \Box T_i$ and thus for all $x \in K$ $x \Vdash T_i$ and for all $x \in K$ $x \Vdash' \Box_{q_i} A_i$. If $a \nVdash T_i$, then $a \Vdash \neg T_i$, $\ a \Vdash \Box\neg T_i$, for all $x \in K$ $x \Vdash \neg T_i$ and for all $x \in K$ $x \Vdash' \neg\Box_{q_i} A_i$.

Let then for some $x \in K$ $\ x \Vdash' \Box_{q_i} A_i$. Then regardless to $x = a$ or $x \ne a$ we have $x \Vdash T_i \to A_i^t$, $x \Vdash' \Box_{q_i} A_i \to A_i$, and $x \Vdash' A_i$.

To complete the proof of Lemma 2.6 assume that $\mathcal{B}^- \nvdash A$. Take the set of all subformulas of A as an adequate set X, get an X-countermodel for A and then extend this X-model to a countermodel of A by putting all atomic and q-atomic formulas not from X to be false at each node. Lemma 2.6 is proved. ■

2.9 Corollary. $\mathcal{B}^-$ enjoys the finite model property and thus $\mathcal{B}^-$ is decidable.

2.10 Remark. The proof of the Kripke style completeness for $\mathcal{B}^-$ provides effectively a finite set of "possible countermodels" to A. This gives an upper bound on the computational time for a decision procedure of order 2^{cn}, with n the length of the formula, and c fixed.

We proceed with the proof of Theorem 2.5, direction ($\Longrightarrow$). After the correctness of $\mathcal{B}^-$ we have only to verify that if A is obtained by **R3** rule from $\Box A$, which is valid in all $\Box A$-sound models, then A is valid in all A-sound models. Suppose A is false at some node of an A-sound model $\mathcal{K} = (K, \prec, \Vdash)$ with the root node a. We construct a new model $\mathcal{K}' = (K', \prec', \Vdash')$ by adding the new root node b (i.e. $K' = K \cup \{b\}$, $b \prec' x \in K$ and $\prec' = \prec$ on K), and defining for every atomic and q-atomic formula Q

$$b \Vdash' Q \iff Q \in SbA \text{ and } a \Vdash Q.$$

Now $\prec'$ has only one extension from atomic and q-atomic formulas to all modal formulas, satisfying forcing conditions 1 and 2; we have to check now that this extension satisfies 3 and 4 as well. Everything is clear with the stability condition. Note that $\prec'$ coincides with $\prec$ on K, and thus q-reflexivity holds at each node of $\mathcal{K}'$ but, may be, the root node b. As $\mathbf{H}(\Box A)$ is valid at the root node a of the model $\mathcal{K}$, by an easy induction one can show that for each $B \in SbA$

$$b \Vdash' B \iff a \Vdash B.$$

The q-reflexivity condition for $\mathcal{K}'$ at the root node b is now obvious: if $\Box_p B \in SbA$, then it works at b the same way as at a, if $\Box_p B \notin SbA$, then $\Box_p B$ is false at b by the definition.

Case ($\Longleftarrow$). Suppose $\mathcal{B} \nvdash A$ and let N be the cardinality of $Z = \{\Box B \mid \Box B \in SbA\}$. Let $\mathcal{K} = (K, \prec, \Vdash)$ be a countermodel for $\Box^{N+1}A$. Then there is a sequence of nodes $a_0 \prec a_1 \prec \ldots \prec a_{N+1}$ such that $a_i \Vdash \neg\Box^{N+1-i}A$, with $0 \leq i \leq N+1$. None of the formulas $\Box B \rightarrow B$ ($\Box B \in Z$) can be false at two different nodes a_i and a_j. So, by the pigeonhole principle there is an i ($0 \leq i \leq N+1$) such that $a_i \Vdash \mathbf{H}(A)$. As soon as $a_i \prec a_{N+1}$ or $a_i = a_{N+1}$ the restriction $\mathcal{K}'$ of the model $\mathcal{K}$ to the set of its nodes $\{b \mid a_i \prec b \text{ or } a_i = b\}$ is an A-sound countermodel for A.
■

2.11 Corollary. $\mathcal{B}$ is decidable.

Proof. As we could see above

$$\mathcal{B} \vdash A \quad \text{iff} \quad \mathcal{B}^- \vdash \Box^{N+1}A,$$

where N is the cardinality of $Z = \{\Box B \mid \Box B \in SbA\}$. As $\Box^{N+1}A$ is not longer than the double length of A the upper bound of the complexity of $\mathcal{B}$ is also of order 2^{cn}, with n the length of the formula, and c fixed.
■

2.3 The arithmetical completeness of $\mathcal{B}$

2.12 Theorem.

$$\mathcal{B} \vdash A \iff \text{for every interpretation } * \quad \mathtt{PA} \vdash A^*$$

Proof. Correctness (i.e. the case $\Longrightarrow$). Induction on a proof of A in $\mathcal{B}$. The cases of axioms **A1-A3** are treated in [7] ($\mathtt{GL}$-correctness for standard proof predicates), **A5, A6** are trivial because both $\mathit{Prf}(p^*, B^*)$ and $\neg\mathit{Prf}(p^*, B^*)$ are recursive. Let us take the axiom **A4**: $\Box_p B \rightarrow B$. Consider two cases: if $\mathit{Prf}(p^*, B^*)$ is true, then B^* is in fact provable and $\mathtt{PA} \vdash (\Box_p B \rightarrow B)^*$; if $\mathit{Prf}(p^*, B^*)$ is false, then $(\Box_p B)^*$ is a false recursive formula and thus $\mathtt{PA} \vdash (\neg\Box_p B)^*$ and again $\mathtt{PA} \vdash (\Box_p B \rightarrow B)^*$.

($\Longleftarrow$). Let $\mathcal{B} \not\vdash A$. Take $X = SbA$ as a finite adequate set. Let also $\mathcal{K} = (K, \prec, \Vdash)$ be an A-sound SbA-model such that A is false at some node of $\mathcal{K}$. We assume that $K = \{1, \ldots, n\}$ and 1 is the root node. As in [3] we define a new model $\mathcal{K}'$ by adding a node 0 to K, putting $0 \prec i$ $(1 \leq i \leq n)$, and defining $0 \Vdash Q$ iff $1 \Vdash Q$ for every atomic and q-atomic formula $Q \in SbA$.

2.13 Lemma. $\mathcal{K}'$ is an A-sound SbA-model.

Proof. Like in the proof of Theorem 2.5, direction "$\Longleftarrow$" (correctness). It suffices to show that for every $G \in SbA \quad 0 \Vdash G$ iff $1 \Vdash G$. Induction on G. The basis, i.e. the cases of atomic and q-atomic formulas, follows from the definition of $\mathcal{K}'$. The case of Boolean connectives is trivial, the step corresponding to $\Box$ operator holds because $1 \Vdash \Box D \rightarrow D$ for every formula $\Box D \in SbA$.
■

Now for the model $\mathcal{K}'$ and for the usual Gödel proof predicate $\mathit{Proof}(x, y)$ we define a Solovay function $h(t)$:

- $h(0) = 0$;
- if $h(m) = i$ and $\mathit{Proof}(m, \ulcorner l \neq j \urcorner)$ for some $j \succ i$, set $h(m+1) = j$; otherwise put $h(m+1) = h(m)$,

where "$l = j$" is a natural arithmetical formula for "j is a limit of $h(t)$".

We also assume that the following Solovay lemma holds:

2.14 Lemma. ([3])

1. $\mathtt{PA} \vdash$ "$0 \leq l \leq n$";
2. "$l = 0$" is true, but each of the theories $\mathtt{PA}$ + "$l = i$" is consistent for $i = 0, 1 \ldots, n$;
3. $\mathtt{PA}$ + "$l = i$" $\vdash \mathit{Provable}(\ulcorner$"$l \neq i$"$\urcorner)$, $\quad i = 1, 2, \ldots, n$;
4. $\mathtt{PA}$ + "$l = i$" $\vdash \neg\mathit{Provable}(\ulcorner$"$l \neq j$"$\urcorner)$, $\quad i = 0, 2, \ldots, n,\ i \prec j$;
5. $\mathtt{PA}$ + "$l = i$" $\vdash \neg\mathit{Provable}(\ulcorner$"$l = j$"$\urcorner)$, $\quad i = 1, 2, \ldots, n,\ i \not\prec j$.

In order to define an arithmetical interpretation $* = (\mathit{Prf}, \varphi)$ we first introduce φ for proof and sentence variables occurring in SbA:

$$\varphi(S_i) := [\bigvee_{j \Vdash S_i} \text{“}l = j\text{”}] \wedge i = i\text{[4]},$$

$$\varphi(p_j) := 2j.$$

Without loss of generality we assume that

$$T \; = \; \{ \; \Box_{p_i} A_{i,j} \mid 0 \leq i \leq m, \; 0 \leq j \leq J_i \; \}$$

is the set of all q-atomic formulas from SbA which are valid in the model $\mathcal{K}'$. By the fixed point argument one can find a predicate Prf – and this Prf completes the interpretation $*$ – such that the following *Fixed Point Equation* (FPE) is provable in $\mathtt{PA}$:

$$\begin{array}{llllll}
\mathit{Prf}(u,v) & \longleftrightarrow & \forall r \leq u \Big[& & & \\
& u = 2r+1 & \to & \mathit{Proof}(r,v) & & \& \\
& u = 2r & \to & [\quad r = 0 & \to \; \bigvee_{j=0}^{J_0} (v = \ulcorner A_{0,j}{}^* \urcorner) & \& \\
& & & & \vdots & \\
& & & \quad\; r = m & \to \; \bigvee_{j=0}^{J_m} (v = \ulcorner A_{m,j}{}^* \urcorner) & \& \\
& & & \quad\; r > m & \to \; v = \ulcorner \forall x_0 (x_0 = x_0) \urcorner &] \Big].
\end{array}$$

It is easy to see now that $\mathit{Prf}(u,v)$ is a primitive recursive formula and that the interpretation $*$ is injective. We will demonstrate that $\mathit{Prf}(u,v)$ is a standard proof predicate for $\mathtt{PA}$ after the proof of the Lemma 2.15 below, but so far we can state that

$$\mathtt{PA} \vdash \forall y [\mathit{Pr}(y) \; \leftrightarrow \; \mathit{Provable}(y) \vee \bigvee_{i=0}^{i=m} \bigvee_{j=0}^{j=J_i} y = \ulcorner A_{i,j}{}^* \urcorner].$$

2.15 Lemma.

Let $G \in SbA$ and $0 \leq k \leq n.$ then

$$k \Vdash G \quad \Longrightarrow \quad \mathtt{PA} \vdash \text{“}l = k\text{”} \to G^*,$$

$$k \nVdash G \quad \Longrightarrow \quad \mathtt{PA} \vdash \text{“}l = k\text{”} \to \neg G^*.$$

Proof. By induction on formulas. Basis in case when G is a sentence variable, or a constant $\bot$ is trivial. Let G be a q-atomic formula. If $k \Vdash G$, then for some $\Box_{p_i} A_{i,j} \in T \;\; G \equiv \Box_{p_i} A_{i,j}$ and $G^* \equiv \mathit{Prf}(2i, \ulcorner A_{i,j}{}^* \urcorner)$, which is a true primitive

[4] A conjunct $i = i$ is added to ensure the injectivity of the arithmetical interpretation $*$ below.

recursive formula according to FPE. Then $\mathtt{PA} \vdash G^*$ and $\mathtt{PA} \vdash$ "$l = k$" $\to G^*$. Let $k \nVdash G$ and $G \equiv \Box_{p_i} H$ for some H. Consider two cases. If $i \leq m$, then $H \not\equiv A_{i,j}$ for every $j = 0, 1, \ldots, J_i$. By the injectivity of $*$ $\quad H^* \not\equiv A_{i,j}{}^*$, thus $\ulcorner H^* \urcorner \neq \ulcorner A_{i,j}{}^* \urcorner$ for $j = 0, 1, \ldots, J_i$ and $\mathit{Prf}(2i, \ulcorner H^* \urcorner)$ is false. But $G^* \equiv \mathit{Prf}(2i, \ulcorner H^* \urcorner)$, thus $\mathtt{PA} \vdash \neg G^*$ and $\mathtt{PA} \vdash$ "$l = k$" $\to \neg G^*$. Let $i > m$. Then $G^* \equiv \mathit{Prf}(2i, \ulcorner H^* \urcorner)$ and thus G^* is false by FPE. Indeed, $\mathit{Prf}(u, v)$ is functional for each $u > 2m$, $\mathit{Prf}(2i, \ulcorner \forall x_0 (x_0 = x_0) \urcorner)$ is true and $\forall x_0 (x_0 = x_0) \not\equiv H^*$ for any modal formula H. Again we have $\mathtt{PA} \vdash \neg G^*$ and $\mathtt{PA} \vdash$ "$l = k$" $\to \neg G^*$.

Now we proceed with different inductions on formulas, first for $k > 0$, and then for $k = 0$.

Let $1 \leq k \leq n$. The induction step in case of "$\to$" is straightforward (cf.[3]). The induction step in case $G \equiv \Box H$: we proceed with the standard Solovay argument.

If $k \Vdash \Box H$, then

$$\text{for all } \; j \succ k \;\; j \Vdash H,$$

$$\text{for all } \; j \succ k \;\; \mathtt{PA} \vdash \text{“}l = j\text{”} \to H^*,$$

$$\mathtt{PA} \vdash \bigvee_{k \prec j} \text{“}l = k\text{”} \to H^*,$$

$$\mathtt{PA} \vdash \mathit{Provable}(\bigvee_{k \prec j} \text{“}l = k\text{”}) \to \mathit{Provable}(H^*), \qquad\qquad \dagger$$

$$\text{by 2.14 (4)} \;\; \mathtt{PA} \vdash \text{“}l = k\text{”} \to \bigwedge_{k \not\prec j} \mathit{Provable}(\ulcorner \text{“}l \neq j\text{”} \urcorner),$$

$$\mathtt{PA} \vdash \text{“}l = k\text{”} \to \mathit{Provable}(\ulcorner \bigwedge_{k \not\prec j} \text{“}l \neq j\text{”} \urcorner) \text{ (by commutting } \mathit{Provable}(\cdot) \text{ and } \bigwedge),$$

$$\text{by 2.14 (1)} \;\; \mathtt{PA} \vdash \mathit{Provable}(\ulcorner \bigvee_{0 \leq j \leq n} \text{“}l = j\text{”} \urcorner),$$

$$\mathtt{PA} \vdash \mathit{Provable}(\ulcorner \bigwedge_{k \neq j} \text{“}l \neq j\text{”} \to \bigvee_{k = j, k \prec j} \text{“}l = j\text{”} \urcorner),$$

$$\text{by 2.14 (3)} \;\; \mathtt{PA} \vdash \text{“}l = k\text{”} \to \mathit{Provable}(\ulcorner \text{“}l \neq k\text{”} \urcorner),$$

$$\text{finally, } \; \mathtt{PA} \vdash \text{“}l = k\text{”} \to \mathit{Provable}(\ulcorner \bigvee_{k \prec j} \text{“}l = j\text{”} \urcorner),$$

$$\text{and by } \dagger \; \mathtt{PA} \vdash \; \text{“}l = k\text{”} \to \mathit{Provable}(\ulcorner H^* \urcorner),$$

$$\text{and as } \mathtt{PA} \vdash \forall y (\mathit{Provable}(y) \to \mathit{Pr}(y))$$

$$\text{we are done: } \mathtt{PA} \vdash \; \text{“}l = k\text{”} \to G^*.$$

If $k \nVdash \Box H$, then

$$\text{for some } \ j \succ k \ \ j \nVdash H,$$

$$\text{by the induction hypothesis } \mathtt{PA} \vdash \text{“}l = k\text{”} \to \neg H^*,$$

$$\mathtt{PA} \vdash H^* \to \text{“}l \neq j\text{”},$$

$$\mathtt{PA} \vdash \mathit{Provable}(\ulcorner H^* \urcorner) \to \mathit{Provable}(\ulcorner \text{“}l \neq j\text{”} \urcorner),$$

$$\mathtt{PA} \vdash \neg \mathit{Provable}(\ulcorner \text{“}l \neq j\text{”} \urcorner) \to \neg \mathit{Provable}(\ulcorner H^* \urcorner),$$

$$\text{but by 2.14 (4)} \ \ \mathtt{PA} \vdash \text{“}l = k\text{”} \to \neg \mathit{Provable}(\ulcorner \text{“}l \neq j\text{”} \urcorner),$$

$$\text{thus } \mathtt{PA} \vdash \ \text{“}l = k\text{”} \to \neg \mathit{Provable}(\ulcorner H^* \urcorner).$$

But $Pr(y)$ may differ from $\mathit{Provable}(y)$ only on the Gödel numbers of the q-atomic formulas from the fixed finite set T; every formula from T is valid in each node of the model $\mathcal{K}'$, thus $H \notin T$, and

$$\mathtt{PA} \vdash \ Pr(\ulcorner H^* \urcorner) \to \mathit{Provable}(\ulcorner H^* \urcorner).$$

Finally, we have got the desired $\mathtt{PA} \vdash$ “$l = k$” $\to \neg Pr(\ulcorner H^* \urcorner)$, i.e.

$$\mathtt{PA} \vdash \ \text{“}l = k\text{”} \to \neg G^*.$$

Let now $k = 0$. Again, the basis of the induction on formulas is done, the case of $\to$ is trivial. Let $G = \Box H$. If $0 \Vdash \Box H$, then for all $j = 1, \ldots, n \ \ j \Vdash H$, by the previous induction

$$\text{for all } \ j = 1, \ldots, n \ \ \mathtt{PA} \vdash \text{“}l = j\text{”} \to H^*,$$

$$\mathtt{PA} \vdash \bigvee_{1 \leq j \leq n} \text{“}l = j\text{”} \to H^*.$$

Also $1 \Vdash H \implies 0 \Vdash H$, and by the induction hypothesis $\mathtt{PA} \vdash$ “$l = 0$” $\to H^*$. By 2.14 (1) $\mathtt{PA} \vdash$ “$0 \leq j \leq n$”, and thus $\mathtt{PA} \vdash H^*$, $\mathtt{PA} \vdash \mathit{Provable}(\ulcorner H^* \urcorner)$, and $\mathtt{PA} \vdash$ “$l = 0$” $\to G^*$. If $0 \nVdash \Box H$, then for some $j > 0 \ \ j \nVdash H$, by the previous induction $\mathtt{PA} \vdash$ “$l = j$” $\to \neg H^*$, $\mathtt{PA} \vdash H^* \to$ “$l \neq j$”,

$$\mathtt{PA} \vdash \mathit{Provable}(\ulcorner H^* \urcorner) \to \mathit{Provable}(\ulcorner \text{“}l \neq j\text{”} \urcorner),$$

$$\mathtt{PA} \vdash \neg \mathit{Provable}(\ulcorner \text{“}l \neq j\text{”} \urcorner) \to \neg \mathit{Provable}(\ulcorner H^* \urcorner),$$

$$\text{but by 2.14 (4)} \ \ \mathtt{PA} \vdash \text{“}l = 0\text{”} \to \neg \mathit{Provable}(\ulcorner \text{“}l \neq j\text{”} \urcorner),$$

$$\text{thus } \mathtt{PA} \vdash \ \text{“}l = 0\text{”} \to \neg \mathit{Provable}(\ulcorner H^* \urcorner).$$

The same argument as above shows that

$$\mathtt{PA} \vdash \ Pr(\ulcorner H^* \urcorner) \to \mathit{Provable}(\ulcorner H^* \urcorner),$$

and we again have

$$\mathtt{PA} \vdash \text{“}l = 0\text{”} \to \neg G^*.$$

■

2.16 Corollary. *Prf* is a standard proof predicate.

Proof. We prove even more:

$$\mathtt{PA} \vdash \forall y(Pr(y) \ \leftrightarrow \ Provable(y)).$$

Indeed,

$$\mathtt{PA} \vdash \forall y(Pr(y) \ \to \ Provable(y))$$

because of FPE, and $Pr(y)$ may differ from $Provable(y)$ only when $y = \ulcorner G^* \urcorner$ for some $G \in T$, but every such G is a q-atomic formula, which is true in all nodes of the model $\mathcal{K}'$. By the Lemma 2.15

$$\mathtt{PA} \vdash \ [\bigvee_{k=0}^{k=n} \text{“}l = k\text{”}] \to G^*,$$

but by Lemma 2.14 (1)

$$\mathtt{PA} \vdash \ \bigvee_{k=0}^{k=n} \text{“}l = k\text{”},$$

thus $\mathtt{PA} \vdash \ G^*$, and $\mathtt{PA} \vdash \ Provable(\ulcorner G^* \urcorner)$. Therefore

$$\mathtt{PA} \vdash \ y = \ulcorner G^* \urcorner \to Provable(y),$$

which implies the desired

$$\mathtt{PA} \vdash \forall y(Pr(y) \ \leftrightarrow \ Provable(y)).$$

■

The final steps of the proof of the theorem are standard. As $A \in SbA$ and $k \not\Vdash A$ for some $k \in K$, we have

$$\mathtt{PA} \vdash \text{“}l = k\text{”} \to \neg A^*,$$

$$\mathtt{PA} \vdash A^* \to \neg\text{“}l = k\text{”}.$$

Suppose $\mathtt{PA} \vdash A^*$, then $\mathtt{PA} \vdash \neg\text{“}l = k\text{”}$, which contradicts Lemma 2.14 (2). Theorem 2.12 is thus proved.

■

2.17 Example. Now we are equipped to prove that $\neg\Box_p\Box_p\top$ is false for some standard proof predicate. Consider a countermodel where the only q-atomic formula $\Box_p\top$ holds, and then apply the arithmetical completeness theorem.

2.18 Notes. 1. The set of those modal formulas, that are true under each interpretation, can be axiomatized exactly as in [3] by the system $\mathcal{B}'$ which has as axioms all theorem of $\mathcal{B}$ and all formulas $\Box A \to A$; and as its only deduction rule *modus ponens.*

2. The arithmetical completeness theorem implies, that $\mathcal{B}$ is conservative over the provability logic $\mathtt{GL}$ (with the usual modality $\Box$ only), and both $\mathcal{B}$ and $\mathcal{B}'$ are conservative over the system $\mathcal{P}$ from [1] (with the labeled modalities only).

3. As in [1] one can construct a uniform proof predicate for the systems $\mathcal{B}$ and $\mathcal{B}'$, i.e. such standard predicate *Prf* that $\mathcal{B}$ and $\mathcal{B}'$ are complete with respect to arithmetical interpretations based on this particular *Prf*. But unlike the case of the usual provability logic $\mathtt{GL}$ the system $\mathcal{B}$ doesn't allow uniform interpretation of sentence variables.

3 Logic of functional proof predicates

The language $\mathcal{L}^+$ is able to distinguish some functional proof predicates; for example, the formula $\neg\Box_p\neg\Box_p\bot$ is true for the usual Gödel proof predicate *Proof* (y) (Chapter 4), but it is false for some other functional proof predicates (Example 3.18). So one cannot expect to get one logic which would be complete for every single functional proof predicate. In this chapter we prove that $\mathcal{F}$ is complete in the entire class of functional proof predicates.

3.1 Kripke semantics for $\mathcal{F}$

Nothing unexpected happens in this section, and it is good news; $\mathcal{F}$ is proven to enjoy a natural finite Kripke semantics, although it required some technical maneuvers to incorporate unification into the standard Kripke environment.

3.1 Definition. A model is called **functional** if it satisfies an extra forcing condition:

5. $x \Vdash \neg(\Box_p B \& \Box_p C)$ if B, C are not unifiable, else
$x \Vdash \Box_p B \& \Box_p C \Longrightarrow x \Vdash D \to E$ for every D, E s.t. $D = E \pmod{B = C}$.
(functionality)

3.2 Theorem.

$$\mathcal{F} \vdash A \quad \Longleftrightarrow \quad A \text{ is valid in all functional } A\text{-sound models.}$$

The **proof** of Theorem 3.2 will be completed at the very end of Subsection 3.1. Again, for technical reasons we consider first an auxiliary system $\mathcal{F}^-$, which has the same axioms as $\mathcal{F}$ and rules **R1** and **R2**.

3.3 Lemma.

$$\mathcal{F}^- \vdash A \quad \Longleftrightarrow \quad A \text{ is valid in all functional models.}$$

Proof. Correctness, i.e. direction ($\Longrightarrow$) is trivial, because we built in the definition of a model all we need for the straightforward induction on the proof in $\mathcal{F}^-$, e.g. functionality condition stands for **A7**.

The direction ($\Longleftarrow$). We proceed first with a few fairly standard steps, like the canonical model for $\mathcal{F}^-$, and then we'll see how filtration meets unification. A set α of modal formulas is $\mathcal{F}^-$**-consistent** if for no $A_1, \ldots, A_n \in \alpha \quad \mathcal{F}^- \vdash \neg(A_1 \& \ldots \& A_n)$; α is a maximal set if either $B \in \alpha$ or $\neg B \in \alpha$ for every modal formula B. It is clear that each consistent set is included into some maximal consistent set. If we then define W to be the class of all maximal $\mathcal{F}^-$-consistent sets, accessibility relation R on W as

$$\alpha R \beta \quad \text{ iff } \quad \text{for every modal formula } F \ (\Box F \in \alpha \Longrightarrow F \in \beta),$$

and put

$$\alpha \Vdash F \quad \text{iff} \quad F \in \alpha,$$

we get the **canonical model** $\mathcal{CM} = (W, R, \Vdash)$. The following standard lemma holds:

3.4 Lemma. Each node α of $\mathcal{CM}$ contains all theorems of $\mathcal{F}^-$ and is closed under *modus ponens.*

For a proof one may consult, for example, [10]. As a result, the forcing conditions 1-5 hold in $(W, R, \Vdash)$.

3.5 Corollary.

$$\mathcal{F}^- \vdash A \quad \Longleftrightarrow \quad \mathcal{CM} \Vdash A.$$

Proof. By the previous lemma all theorems of $\mathcal{F}^-$ hold in $\mathcal{CM}$. If $\mathcal{F}^- \nvdash A$, then the set $\{\neg A\}$ is consistent; take a maximal consistent set α containing $\neg A$. According to 3.5 $\alpha \nVdash A$.
■

So $\mathcal{CM}$ is a countermodel[5] for any A not-theorem of $\mathcal{F}^-$, but $\mathcal{CM}$ has continuum many nodes. Now for a given A such that $\mathcal{F}^- \nvdash A$ we construct a finite tree countermodel $\mathcal{K}_T = (K, \prec, \Vdash_T)$ by stages (cf.[9]). K will consist of finite R-increasing sequences in $\mathcal{CM}$, and $\prec$ will be the usual strict ordering by extension of finite sequences.

[5] although $\mathcal{CM}$ is not a model in our *finite tree* sense

- Stage 0. Pick (α_0) such that $\alpha_0 \in W$ and $\alpha_0 \not\Vdash A$ as the root node of $\mathcal{K}_T$.
- Stage $n+1$. For each sequence $(\alpha_0, \ldots, \alpha_n)$ already in K, look at $\{\Box B \in SbA \mid \alpha_n \not\Vdash \Box B\}$. If the set is empty, do not extend $(\alpha_0, \ldots, \alpha_n)$. Otherwise, for each such $\Box B$, choose a node $\beta \in W$ such that $\alpha_n R \beta$, $\beta \Vdash \Box B$ and $\beta \not\Vdash B$. Such β should exist: $\alpha_n \Vdash \Box(\Box B \to B) \to \Box B$ (this is axiom **A3**), thus $\alpha_n \not\Vdash \Box(\Box B \to B)$ and there is β, $\alpha_n R \beta$, where $\beta \not\Vdash \Box B \to B$, i.e. $\beta \not\Vdash B$ and $\beta \Vdash \Box B$. Add $(\alpha_0, \ldots, \alpha_n, \beta)$ to K and define

  $$(\alpha_0, \ldots, \alpha_n, \beta) \Vdash_T D \quad \text{iff} \quad \beta \Vdash D$$

  for any modal formula D.

That this is a tree with the root (α_0) is obvious. Finiteness follows from König's Lemma: the tree is finitely branching because branches are correlated with elements of the finite set SbA and there are no infinite paths because the succession from $(\alpha_0, \ldots, \alpha_n)$ to $(\alpha_0, \ldots, \alpha_n, \alpha_{n+1})$ results in at least one additional sentence $\Box B \in SbA$ being forced by α_{n+1}. Forcing conditions 1,4, and 5 hold because they regulate a pointwise behaviour of the forcing relation, which by definition is similar for $\Vdash_T$ and $\Vdash$. Condition 2 is fulfilled by the special choice of nodes of W extending sequences in K. The stability condition 3 is guaranteed by the fact that all nodes β of $\mathcal{CM}$ occurring in the sequences in K are R-accessible from α_0; if $\alpha_0 \Vdash \Box_p B$, then by $\alpha_0 \Vdash \Box_p B \to \Box\Box_p B$ (axiom **A5**) we have $\alpha_0 \Vdash \Box\Box_p B$ and $\beta \Vdash \Box_p B$. If $\alpha_0 \Vdash \neg\Box_p B$, then use **A6**. Now, $\mathcal{K}_T$ is a functional model, $\alpha_0 \not\Vdash A$, and thus $(\alpha_0) \not\Vdash_T A$ and this completes the proof of Lemma 3.3.
■

3.6 Corollary. $\mathcal{F}^-$ enjoys a finite model property.

It doesn't give us, however, a decision algorithm for $\mathcal{F}^-$ directly. A forcing relation in a model is not inductively defined, and even a finite countermodel for A requires *a priori* a check of the functionality axiom **A7**, which involves arbitrary formulas D, E, not only subformulas of A. We still have to show how to extend a partial forcing relation defined on the subformulas of A to a real forcing relation on all formulas.

For a modal formula A and a frame $(K, \prec)$ let $x \Vdash_0 B$ be a relation defined on the nodes $x \in K$ and on $B \in SbA$. We assume that $\Vdash_0$ satisfies forcing conditions 1-4; as none of these conditions refers to any formula other than from SbA, there is a clear algorithm which for a given A, $(K, \prec)$, and $\Vdash_0$ decides whether the forcing conditions 1-4 are fulfilled. Below, while dealing with stable (partial) forcing relations we will write $\Vdash Q$ instead of $x \Vdash Q$ for a Boolean combination Q of q-atomic formulas.

We describe now decidable necessary and sufficient conditions for A, $(K, \prec)$, and $\Vdash_0$, which in addition to 1-5 guarantee that $\Vdash_0$ can be extended to a real forcing relation $\Vdash$ on the frame $(K, \prec)$. Let $\mathcal{V}$ be the following algorithm which starts with a frame $(K, \prec)$, a relation $\Vdash_0$ defined for $x \in K,\ G \in SbA$ satisfying forcing conditions 1-4 for all subformulas of A.

- Step 0: $Y_0 := SbA,\ \tau_0 :=$ the empty substitution ϵ,
- Step $i+1$:
  1. Pick any pair $\Box_p B, \Box_p C \in Y_i$ such that $B \not\equiv C$, $\Vdash_i \Box_p B$ and $\Vdash_i \Box_p C$; if there in no such a pair, terminate with the announcement that $\tau = \tau_0\tau_1, \ldots, \tau_i$.
  2. Run the unification algorithm $\mathbf{U}$ to obtain the standard mgu τ_{BC} of B and C, put $\tau_{i+1} := \tau_{BC}$; if B, C are not unifiable, then $\mathcal{V}$ fails.
  3. Check whether there are $x \in K,\ D_1, D_2 \in Y_i$ such that $x \Vdash_i D_1,\ x \nVdash_i D_2$, but $D_1\tau_{i+1} \equiv D_2\tau_{i+1}$. If *Yes*, then $\mathcal{V}$ fails; if *No*, then $Y_{i+1} := Y_i\tau_{i+1}$, and for every $x \in K, D \in Y_i$ put
  $$x \Vdash_{i+1} D\tau_{i+1} \qquad \text{iff} \qquad x \Vdash_i D.$$
  Note that Y_{i+1} remains closed under subformulas.
  4. Check stability and q-reflexivity properties of $\Vdash_{i+1}$ on Y_{i+1}. If *No*, then $\mathcal{V}$ fails.

Algorithm $\mathcal{V}$ apparently terminates as it reduces the number of variables in Y_i. By induction on i one can demonstrate that Y_i are all closed under subformulas. Indeed, the basis $i = 0$ is an assumption, let now $F \in Y_i\tau_{i+1},\ \tau_{i+1}$ be

$$(T_1 \leftarrow A_1, \ldots T_n \leftarrow A_n, q_1 \leftarrow r_1, \ldots q_m \leftarrow r_m),$$

and $G \in SbF$. Then $F \equiv F'\tau_{i+1}$ for some $F' \in Y_i$, and by an easy induction on a formula F' one can show that either $G \equiv G'\tau_{i+1}$ for some $G' \in SbF'$, or $G \in SbA_j$ for some $j = 1, \ldots, n$. In the first of these cases trivially $G \in Y_i\tau_{i+1}$, in the second $G \in Y_i$, no variable from G is in $Dom(\tau_{i+1})$, $G\tau_{i+1} \equiv G$, and again $G \in Y_i\tau_{i+1}$. It is also clear that $\Vdash_i$ satisfies forcing conditions 1-4 for formulas from Y_i, and for each $B \in SbA,\ x \in K,\ i = 0, 1, \ldots$

$$x \Vdash_0 B \iff x \Vdash_i B\tau_o\tau_1 \cdots \tau_i.$$

We claim that $\Vdash_i,\ i = 0, 1, \ldots$, are all subsets of a desired forcing relation $\Vdash$, provided that the later exists.

3.7 Lemma. If $(K, \prec, \Vdash)$ is a functional model and for all $x \in K,\ B \in SbA$

$$x \Vdash B \iff x \Vdash_0 B,$$

then for every $i = 0, 1, \ldots,\ D \in Y_i$

$$x \Vdash_i D \iff x \Vdash D.$$

Proof. Induction on $i = 0, 1, \ldots$. The basis is the condition of the lemma. The transition from $\Vdash_i$ to $\Vdash_{i+1}$ goes with $\tau_{i+1} = \tau_{BC}$ for some $\Box_p B, \Box_p C \in Y_i$ such that $\Vdash_i \Box_p B$ and $\Vdash_i \Box_p C$. By the induction hypothesis B, C are necessarily unifiable as the functionality property holds for $\Vdash_i$ and $\Box_p B, \Box_p C$. Let $D \in Y_i$. If $x \Vdash_i D$, then by the indiction hypothesis $x \Vdash \Box_p B$, $x \Vdash \Box_p C$, and $x \Vdash D$, and by the local functionality principle for $\Vdash$ we have also $x \Vdash D\tau_{i+1}$. On the other hand, $x \Vdash_{i+1} D\tau_{i+1}$ by the definition of $\Vdash_{i+1}$. If $x \nVdash_i D$, then by the indiction hypothesis $x \Vdash \Box_p B$, $x \Vdash \Box_p C$, and $x \nVdash D$, and by the local functionality for $\Vdash$ we have $x \nVdash D\tau_{i+1}$. Also $x \nVdash_{i+1} D\tau_{i+1}$ by the definition of $\Vdash_{i+1}$.
■

3.8 Corollary. If $\mathcal{V}$ fails, then there is no model with a desired forcing relation extending $\Vdash_0$. Indeed, let $\mathcal{V}$ fail when it meets a pair of formulas $\Box_p B, \Box_p C \in Y_i$ such that $\Vdash_i \Box_p B$ and $\Vdash_i \Box_p C$, but B and C are not unifiable. By the previous lemma there should be $\Vdash \Box_p B \& \Box_p C$, but by the functionality of the forcing relation $\Vdash \neg(\Box_p B \& \Box_p C)$.

If $\mathcal{V}$ fails when it gets $D_1, D_2 \in Y_i$ such that $x \Vdash_i D_1$, $x \nVdash_i D_2$, but $D_1\tau_{i+1} \equiv D_2\tau_{i+1}$ ($\tau_{i+1} = \tau_{BC}$ for some pair B, C such that $x \Vdash_i \Box_p B$, $x \Vdash_i \Box_p C$), then by the induction hypothesis $x \Vdash D_1$, $x \nVdash D_2$, $x \Vdash \Box_p B \& \Box_p C$ and by the local functionality of $\Vdash$

$$x \Vdash D_1 \iff x \Vdash D_1\tau_{i+1},$$

$$x \Vdash D_2 \iff x \Vdash D_2\tau_{i+1}.$$

Contradiction.

The failure of $\mathcal{V}$ because of the violation of stability or q-reflexivity of $\Vdash_i$ would mean directly that the same property is violated for $\Vdash$.

Now we prove that if $\mathcal{V}$ succeeds, then $\Vdash_0$ can be extended to a real forcing relation $\Vdash$ on $(K, \prec)$. Let $\mathcal{V}$ succeeds and terminates after k steps. Then Y_k is **functional with respect to** $\Vdash_k$, i.e. for all $\Box_p B, \Box_p C \in Y_k$

$$(\Vdash_k \Box_p B \text{ and } \Vdash_k \Box_p C) \Longrightarrow B \equiv C.$$

Note also that the final value Y_k is $(SbA)\tau$. In fact $(SbA)\tau$ coincides with $Sb(A\tau)$: inclusion $(SbA)\tau \subseteq Sb(A\tau)$ is trivial, and $Sb(A\tau) \subseteq (SbA)\tau$ holds because $A\tau \in (SbA)\tau$, and $(SbA)\tau$ is closed under subformulas.

We first define the forcing relation $\Vdash$ on $(K, \prec)$ for atomic and q-atomic formulas Q by

$$x \Vdash Q \qquad \text{iff} \qquad Q\tau \in (SbA)\tau \text{ and } x \Vdash_k Q\tau$$

for every $x \in K$, and then extend it to all modal formulas by the forcing conditions 1 and 2.

3.9 Lemma. $(K, \prec, \Vdash)$ is a functional model.

Proof. We have to verify that $\Vdash$ is a legitimate forcing relation, i.e. forcing conditions 1-5 are respected. Items 1-2. are fulfilled by the definition: any evaluation of atomic and q-atomic formulas at the nodes of a frame can be extended to all modal formulas inductively respecting Boolean connectives and the standard behaviour of the modality $\Box$ (forcing condition 2.) in a unique way.

Stability: for a q-atomic formula Q if $Q\tau \notin (SbA)\tau$, then $\nVdash Q$ by the definition of $\Vdash$, if $Q\tau \in (SbA)\tau$, then $\Vdash$ is stable on Q as $\Vdash_k$ is stable on $Q\tau$.

q-reflexivity: again for a q-atomic $\Box_p B$ if $(\Box_p B)\tau \notin (SbA)\tau$, then $\nVdash \Box_p B$ by the definition of $\Vdash$ and thus in each $x \in K$ $x \Vdash \Box_p B \to B$. If $(\Box_p B)\tau \in (SbA)\tau$, then $B\tau \in (SbA)\tau$, as $(SbA)\tau$ is closed under subformulas, and by the q-reflexivity of $\Vdash_k$ we obtain

$$x \Vdash \Box_p B \implies x \Vdash_k \Box_p B \implies x \Vdash_k B \implies x \Vdash B.$$

Functionality. If $(\Box_p B)\tau \notin (SbA)\tau$, or $(\Box_p C)\tau \notin (SbA)\tau$, or $\nVdash_k (\Box_p B)\tau$, or $\nVdash_k (\Box_p C)\tau$, then $\nVdash \Box_p B \& \Box_p C$ by the definition of $\Vdash$, and for every D, E and $x \in K$

$$x \Vdash \Box_p B \& \Box_p C \to (D \to E).$$

Let now both $\Vdash_k (\Box_p B)\tau$ and $\Vdash_k (\Box_p C)\tau$. We claim that $\tau_{BC}\tau = \tau$. Indeed, then both $\Box_{p\tau} B\tau, \Box_{p\tau} C\tau$ are in $(SbA)\tau$, and as $(SbA)\tau$ is functional with respect to $\Vdash_k$, $B\tau \equiv C\tau$. According to the properties of mgu τ_{BC} there exists a substitution λ such that $\tau = \tau_{BC}\lambda$. Then with the use of Lemma 1.1 we obtain

$$\tau_{BC}\tau = \tau_{BC}(\tau_{BC}\lambda) = (\tau_{BC}\tau_{BC})\lambda = \tau_{BC}\lambda = \tau.$$

We check the functionality forcing condition first for the atomic and q-atomic formulas. Let D, E be atomic or q-atomic and $D = E \pmod{B = C}$, i.e. $D\tau_{BC} \equiv E\tau_{BC}$, then $D\tau \equiv E\tau$ and thus $x \Vdash D \iff x \Vdash E$. Now we extend the functionality property on arbitrary modal formulas D, E by an easy induction on D, E. The basis, i.e. the case of atomic or q-atomic D, E has just been checked. The case of Boolean connectives is pointwise and thus trivial. We proceed with the $\Box$ case. Let for all $x \in K$

$$x \Vdash (\Box_p B \& \Box_p C) \to (D \to E).$$

Then on the basis of the forcing conditions 1-2, we may infer that for all $x \in K$

$$x \Vdash \Box(\Box_p B \& \Box_p C) \to (\Box D \to \Box E),$$

but the stability of $\Vdash$ gives

$$x \Vdash (\Box_p B \& \Box_p C) \implies x \Vdash \Box(\Box_p B \& \Box_p C),$$

and we are done.

■

3.10 Lemma. $(K, \prec, \Vdash)$ is a countermodel for A.

Proof. As $\Vdash$ is a forcing relation extending $\Vdash_0$, and $(\alpha_0) \nVdash_0 A$ we conclude that $(\alpha_0) \nVdash A$.

■

3.11 Corollary. $\mathcal{F}^-$ is decidable.

Proof. Now together with the finite model property for $\mathcal{F}^-$ we have an algorithm, which for a given finite frame and a given partial forcing relation defined on the nodes of this frame and on the subformulas of a given formula decides whether there exists a model extending this partial forcing relation. It provides a recursive enumeration of all finite models and thus all non-theorems of $\mathcal{F}^-$; by the standard Post argument $\mathcal{F}^-$ is decidable. Again, the proof of the Kripke style completeness for $\mathcal{F}^-$ provides effectively a finite set of "possible countermodels" to A. This gives an upper bound on the computational time for a decision procedure of order 2^{cn}, with n the length of the formula, and c fixed.

■

Now we procced with the proof of Theorem 3.2. Direction ($\Longleftarrow$) is treated exactly as the corresponding part of the proof of Theorem 2.5. Suppose $\mathcal{F} \nvdash A$ and let N be the cardinality of $Z = \{\Box B \mid \Box B \in SbA\}$. Then $\mathcal{F}^- \nvdash \Box^{N+1}A$ and we take a countermodel $\mathcal{K} = (K, \prec, \Vdash)$ to the formula $\Box^{N+1}A$. Then there is a sequence of nodes $a_0 \prec a_1 \prec \ldots \prec a_{N+1}$ such that $a_i \Vdash \neg\Box^{N+1-i}A$, with $0 \le i \le N+1$, in particular $a_{N+1} \nVdash A$. None of the formulas $\Box B \to B$ ($\Box B \in Z$) can be false at two different nodes a_i and a_j. So, by the pigeonhole principle there is an i ($0 \le i \le N+1$) such that

$$a_i \Vdash \mathbf{H}(A).$$

As soon as $a_i \prec a_{N+1}$ or $a_i = a_{N+1}$ the restriction $\mathcal{K}'$ of the model $\mathcal{K}$ to the set of its nodes $\{b \mid a_i \prec b \text{ or } a_i = b\}$ is the desired model.

Direction ($\Longrightarrow$) of Theorem 3.2. We'll get this *correctness* statement for granted after the proof of the arithmetical completeness of $\mathcal{F}$ as an easy corollary of the arithmetical correctness of $\mathcal{F}$ and the fact that for a given countermodel for A one may construct an arithmetical interpretation under which A is not provable in $\mathtt{PA}$ (Theorem 3.14). It is difficult however not to try the machinery of Kripke models for $\mathcal{F}$ already developed above. So here is a draft proof of the Kripke model correctness of $\mathcal{F}$.

Induction on the derivability in $\mathcal{F}$ as in Theorem 2.5, direction ($\Longrightarrow$). The only nontrivial part is the step corresponding to the **R3** rule: let A be obtained from $\Box A$ by **R3**, and suppose there is a countermodel $\mathcal{K} = (K, \prec, \Vdash)$ for A with $\mathbf{H}(A)$ valid at its root node a. Add a new root node b to $(K, \prec)$ to get a new frame $(K', \prec')$, i.e. $K' := K \cup \{b\}$, $\prec' := \prec \cup \{b \prec' x \mid x \in K\}$. Also define a partial forcing relation $\Vdash'$ for SbA on $\mathcal{K}'$ by

$$b \Vdash' B \Longleftrightarrow a \Vdash B \quad \text{and} \quad x \Vdash' B \Longleftrightarrow x \Vdash B$$

for all $x \in K$, $B \in SbA$.

Note that $\Vdash'$ satisfies forcing conditions 1-4 on $\mathcal{K}'$ and SbA, which is guaranteed by $a \Vdash' \mathbf{H}(A)$. Indeed, 1, 3, and 4 are due to the definitions. Condition 2 holds for all nodes of K because $\mathcal{K}$ is a model. Let now $b \Vdash' \Box B$ for some $\Box B \in SbA$. Then $a \Vdash' \Box B$ by the definition of $\Vdash'$. Now $a \Vdash' B$ because $\mathbf{H}(A)$ is valid at a, for all other $x \in K$ $x \Vdash' B$ because 2 holds for a. Thus $x \Vdash' B$ for all $x \in K$, and 2 holds for b as well.

If now the partial forcing relation $\Vdash'$ on $(K', \prec')$ can be extended to a real one, then we are done. Indeed, then we would have

$$a \nVdash A \implies a \nVdash' A \implies b \nVdash' \Box A \implies b \Vdash' \Box A \to A,$$

and $b \Vdash' \mathbf{H}(\Box A)$, as $\mathbf{H}(\Box A)$ is equivalent in a propositional logic to $(\Box A \to A) \& \mathbf{H}(A)$ and $b \Vdash' \mathbf{H}(A)$. So, it suffices now to show that the algorithm $\mathcal{V}$ will succeed on $\mathcal{K}'$, SbA.

Run $\mathcal{V}$ on $\mathcal{K}'$, SbA, and concurrently repeat the same steps of $\mathcal{V}$ on $\mathcal{K}, SbA$. We first notice that if a run of $\mathcal{V}$ on $\mathcal{K}', SbA$ produces sequences

$$Y_0, Y_1, \dots, \qquad \tau_0, \tau_1, \dots, \qquad \text{and} \qquad \Vdash'_0, \Vdash'_1, \dots,$$

then these sentences fit for the relevant run on $\mathcal{K}', SbA$ as well: just $\mathcal{V}$ has to check less nodes on $\mathcal{K}'$ than on $\mathcal{K}$. As $\mathcal{V}$ saves the truth values of all formulas from SbA pointwise, we have for all $B \in SbA$, $i = 0, 1, \dots$

$$b \Vdash'_i B\tau_0\tau_1 \dots \tau_i \iff a \Vdash'_i B\tau_0\tau_1 \dots \tau_i.$$

Then we prove that any failure of $\mathcal{V}$ on $\mathcal{K}'$, SbA would mean a relevant failure of $\mathcal{V}$ on $\mathcal{K}, SbA$, which is impossible as soon as $\mathcal{K}$ is a functional model. Indeed, there might be 3 sorts of reasons for $\mathcal{V}$ to fail on $\mathcal{K}'$, SbA.

1. $\mathcal{V}$ meets a pair $\Box_p B, \Box_p C \in Y_i$ such that $\Vdash'_i \Box_p B, \Vdash'_i \Box_p C$, but B and C are not unifiable. But the same pair would stop $\mathcal{V}$ on $\mathcal{K}, SbA$ as $\Vdash_i$ is a subrelation of $\Vdash'_i$.
2. There are $D_1, D_2 \in Y_i$ such that for some $x \in K'$ $x \Vdash'_i D_1$, $x \nVdash'_i D_2$, but $D_1\tau_{i+1} \equiv D_2\tau_{i+1}$. The same D_1, D_2 would stop $\mathcal{V}$ on $\mathcal{K}$; if $x = b$, then $x = a$ would fit as well.

3. Stability or q-reflexivity of Y_i is violated. Again if b is involved in this violation, then a could play the same role, and it would mean a failure of this property of $\mathcal{V}$ on $\mathcal{K}$ as well.

So $\mathcal{V}$ on $\mathcal{K}'$ should succeed, thus $\mathcal{K}' = (K', \prec', \Vdash')$ can be extended to a model, and we are done with the Kripke correctness of $\mathcal{F}$. This completes the proof of Theorem 3.2.

3.12 Theorem. $\mathcal{F}$ is decidable.

Proof. As we could see above

$$\mathcal{F} \vdash A \quad \text{iff} \quad \mathcal{F}^- \vdash \Box^{N+1} A,$$

where N is the cardinality of $Z = \{\Box B \mid \Box B \in SbA\}$. As $\Box^{N+1}A$ is not longer than the double length of A the upper bound of the complexity of $\mathcal{F}$ is also of order 2^{cn}, with n the length of the formula, and c fixed.
■

It turns out that the existence of a substitution τ which makes SbA functional and preserves the truth values of all subformulas of A in a given model is essential not only for Kripke semantics of $\mathcal{F}$, but also will play a key role in the arithmetic completeness proof for $\mathcal{F}$.

3.13 Corollary. If $\mathcal{F} \nvdash A$, then there is a substitution τ and an $A\tau$-sound functional countermodel $\mathcal{K}$ for $A\tau$ such that $(SbA)\tau = Sb(A\tau)$ and $(SbA)\tau$ is functional in $\mathcal{K}$.

Proof. In addition to what we have already proven we should check only that $(\mathbf{H}(A))\tau = \mathbf{H}(A\tau)$, which easily follows from the definitions and the observation that $(SbA)\tau = Sb(A\tau)$.
■

3.2 The arithmetical completeness of $\mathcal{F}$

3.14 Theorem.

$$\mathcal{F} \vdash A \iff \text{for every functional interpretation } * \quad \mathtt{PA} \vdash A^*$$

Proof. Correctness, i.e. the case ($\Longrightarrow$). Induction on a proof of A in $\mathcal{F}$. After Theorem 2.12, direction ($\Longrightarrow$), it only remains to check the correctness of the functionality axiom **A7**.

3.15 Lemma. If $B^* \equiv C^*$, then B, C are unifiable and $* = \tau_{BC} \circ *$.

Proof. Look at the arithmetical interpretation $*$ as at a substitution in a language which combines $\mathcal{L}^+$ and $\mathtt{PA}$ languages. Then $*$ is a unifier of B, C *modulo* 2 minor details both being irrelevant to our case.

1. $*$ is defined on all variables of $\mathcal{L}^+$, not only on the finite subset of them. Well, we'd better just ignore this.
2. $*$ transforms the occurrences of modalities into proof and provability formulas. To deal with this we have just to teach the Unification Algorithm to handle *Prf*() and *Pr*() as modalities, i.e. to replace an equation $Prf(p^*, \ulcorner\varphi^*\urcorner) = Prf(q^*, \ulcorner\psi^*\urcorner)$ by $p^* = q^*, \varphi^* = \psi^*$ and an equation $Pr(\ulcorner\varphi^*\urcorner) = Pr(\ulcorner\psi^*\urcorner)$ by $\varphi^* = \psi^*$.

As τ_{BC} is an idempotent mgu of B, C, then for some $\sharp$

$$* = \tau_{BC} \circ \sharp = (\tau_{BC} \circ \tau_{BC}) \circ \sharp = \tau_{BC} \circ (\tau_{BC} \circ \sharp) = \tau_{BC} \circ *.$$

■

Let us check the correctness of the functionality axiom **A7**:

$$\Box_p B \& \Box_p C \to (D \to E) \quad \text{if} \quad D = E \quad (\text{mod } B = C).$$

If $\mathtt{PA} \vdash (\Box_p B \& \Box_p C)^*$, then $\ulcorner B^* \urcorner = \ulcorner C^* \urcorner$ because of the functionality of the proof predicate $Prf(x, y)$. Thus $B^* \equiv C^*$ by the injectivity of the Godel numbering. Now by $D\tau_{BC} \equiv E\tau_{BC}$ and by Lemma 3.15

$$D^* \equiv D\tau_{BC}{}^* \equiv E\tau_{BC}{}^* \equiv E^*,$$

and trivially D^* implies E^*.

($\Longleftarrow$). Let $\mathcal{F} \not\vdash A$. By 3.13 take a substitution τ and a finite $A\tau$-sound functional countermodel $\mathcal{K}$ for $A\tau$ such that $(SbA)\tau$ is functional in $\mathcal{K}$. We assume that $K = \{1, \ldots, n\}$ and 1 is the root node. Again we define a new model $\mathcal{K}'$ by adding a node 0 to K, putting $0 \prec i$ $(1 \leq i \leq n)$, and defining $0 \Vdash B$ iff $1 \Vdash B$ for every formula $B \in SbA\tau$. For our purposes it is sufficient to have $\Vdash$ defined at the nodes of $\mathcal{K}'$ and on $SbA\tau$ only and not to bother about a consistent extension of $\Vdash$ to all modal formulas (that however can be done). It is clear that $\Vdash$ satisfies forcing conditions 1-5 on $SbA\tau$ at the node 0 as well: conditions 1,3,4 and 5 hold pointwise and thus are transmitted to the node 0 from the node 1. Condition 2 holds because $1 \Vdash \mathbf{H}(A\tau)$:

$$0 \Vdash \Box B \implies 1 \Vdash \Box B \implies 1 \Vdash B \implies \forall k \succ 0 \ k \Vdash B.$$

Now for the model $\mathcal{K}'$ and for the usual Gödel proof predicate $Proof(x, y)$ we again define a Solovay function $h(t)$:

- $h(0) = 0$;
- if $h(m) = i$ and $Proof(m, \ulcorner l \neq j \urcorner)$ for some $j \succ i$, set $h(m+1) = j$; otherwise put $h(m+1) = h(m)$,

where “$l = j$” is a natural arithmetical formula for “j is a limit of $h(t)$”.

We also assume that the Solovay Lemma 2.14 holds.

Let

$$T = \{\Box_{p_{m_0}} A_0, \Box_{p_{m_1}} A_1 \ldots, \Box_{p_{m_q}} A_q\}$$

be the list of all q-atomic formulas from $SbA\mathcal{T}$ which are valid in $\mathcal{K}'$.

In order to define an arithmetical interpretation $\sharp = (Prf, \varphi)$ we first introduce φ for proof and sentence variables occurring in $SbA\mathcal{T}$:

$$\varphi(S_i) := [\bigvee_{j \Vdash S_i} \text{“}l = j\text{”}] \wedge i = i,$$

$$\varphi(p_j) := 2j.$$

The following Fixed Point Equation (FPE) is provable in PA:

$$\begin{array}{llllllll}
Prf(u,v) & \longleftrightarrow & \forall r \leq u \Big[& & & & & \\
& u = 2r+1 & \rightarrow & Proof(r,v) & & \& & & \\
& u = 2r & \rightarrow & [\;\; r = m_0 & \rightarrow & v = \ulcorner A_0{}^{\sharp} \urcorner & & \& \\
& & & & \vdots & & & \\
& & & \;\;\;\;\; r = m_q & \rightarrow & v = \ulcorner A_q{}^{\sharp} \urcorner & & \& \\
& & & \;\;\;\;\; \bigwedge_{i=0}^{i=q} r \neq m_i & \rightarrow & v = \ulcorner \forall x_0 (x_0 = x_0) \urcorner & \Big]\Big], &
\end{array}$$

where $A_i{}^{\sharp}$ is the arithmetical translation of A_i under the interpretation $\sharp = (Prf, \varphi)$.

It is easy to see now that $Prf(u,v)$ is primitive recursive and functional, and that $\sharp$ is injective. We will check that $Prf(u,v)$ is a standard proof predicate for PA after Lemma 3.16 below, but so far we can state that

$$\mathsf{PA} \vdash \forall y [Pr(y) \;\leftrightarrow\; Provable(y) \vee \bigvee_{i=0}^{i=q} y = \ulcorner A_i{}^{\sharp} \urcorner].$$

3.16 Lemma. For any modal formula $B \in SbA\mathcal{T}$ and all $k = 0, 1, \ldots, n$

$$k \Vdash B \;\;\Rightarrow\;\; \mathsf{PA} \vdash \text{“}l = k\text{”} \rightarrow B^{\sharp},$$

$$k \nVdash B \;\;\Rightarrow\;\; \mathsf{PA} \vdash \text{“}l = k\text{”} \rightarrow \neg B^{\sharp}.$$

Proof. By induction on formulas. Basis in the case when B is a Boolean constant or a sentence variable is trivial. Let B be a q-atomic formula. If $k \Vdash B$, then for some $B \equiv \Box_{p_{m_i}} A_i \in T$ and $B^{\sharp} \equiv Prf(2m_i, \ulcorner A_i{}^{\sharp} \urcorner)$, which is a true primitive recursive formula according to FPE. Then $\mathtt{PA} \vdash B^\sharp$ and $\mathtt{PA} \vdash$ "$l = k$" $\to B^\sharp$. Let $k \not\Vdash B$ and $B \equiv \Box_{p_i} H$ for some H. Consider two cases:

$i \in \{m_0, \dots, m_q\}$. Then $H \not\equiv A_{m_i}$, $i = 0, \dots, q$, and, by the injectivity of $\sharp$, $H^\sharp \not\equiv {A_{m_i}}^\sharp$, thus $\ulcorner H^\sharp \urcorner \neq \ulcorner {A_{m_i}}^\sharp \urcorner$ and $Prf(2i, \ulcorner H^\sharp \urcorner)$ is false. But $B^\sharp \equiv Prf(2i, \ulcorner H^\sharp \urcorner)$, thus $\mathtt{PA} \vdash \neg B^\sharp$ and $\mathtt{PA} \vdash$ "$l = k$" $\to \neg B^\sharp$.

$i \notin \{m_0, \dots, m_q\}$. Then $B^\sharp \equiv Prf(2i, \ulcorner H^\sharp \urcorner)$ and thus $B^\sharp$ is false according to FPE. Indeed, $Prf(u, v)$ is functional, $Prf(2i, \ulcorner \forall x_0(x_0 = x_0) \urcorner)$ is true and $\forall x_0(x_0 = x_0) \not\equiv H^\sharp$ for any modal formula H.

The induction steps go exactly as in Lemma 2.15.

■

3.17 Corollary. *Prf* is a standard proof predicate.

Proof. Exactly as in Lemma 2.16 we prove that

$$\mathtt{PA} \vdash \forall y Pr(y) \leftrightarrow Provable(y).$$

■

We will finish now the proof of the Theorem 3.14: as $i \not\Vdash A\tau$ for some $i \in K$, we have

$$\mathtt{PA} \vdash \text{"}l = i\text{"} \to \neg A\tau^\sharp,$$

$$\mathtt{PA} \vdash A\tau^\sharp \to \neg\text{"}l = i\text{"}.$$

Suppose $\mathtt{PA} \vdash A\tau^\sharp$, then $\mathtt{PA} \vdash \neg$"$l = i$", which contradicts Lemma 2.14 (2); so $\mathtt{PA} \nvdash A\tau^\sharp$. Define $* := \tau\sharp$ and get the desired $\mathtt{PA} \nvdash A^*$.
■

3.18 Example. Let us show now that $\neg\Box_p\neg\Box_p\bot$ is false for some functional proof predicate. Consider a countermodel to $\neg\Box_p\neg\Box_p\bot$ where $\Box_p\neg\Box_p\bot$ holds but all other q-atomic formulas do not hold, conclude that $\neg\Box_p\neg\Box_p\bot$ is not provable in $\mathcal{F}$ and apply Theorem 3.13.

Notes 2.18 remain valid also for the logic $\mathcal{F}$.

4 Logic of the Gödel proof predicate

Let us return to the example of the formula $\neg\Box_p\neg\Box_p\bot$, which is false under some functional interpretation $*$. In fact this formula is true (and provable in $\mathtt{PA}$) under each interpretation based on the usual Gödel proof predicate $Proof(x, y)$. The reason for it is that the usual Gödel numbering is provably monotone in a sense that the Gödel number of a proof is greater, than the Gödel numbers of formulas from this proof, the Gödel number of a formula is greater, than the Gödel number of any term from this formula, the Gödel number of a numeral $\overline{n}$ is greater, than n itself.

4.1 Definition. A functional interpretation $*$ is called **Gödel interpretation** if it is based on the Gödel proof predicate $\mathit{Proof}(x, y)$.

In this chapter we prove that the logic $\mathcal{M}$ is complete with respect to all Gödel interpretations. A "naive" question arises, what is the usual Gödel proof predicate? There are different styles of Gödel numbering and syntactically many different ways to write $\mathit{Proof}(x, y)$. Fortunately we may skip this question because the language $\mathcal{L}^+$ does not distinguish these proof predicates; moreover, the logic $\mathcal{M}$ is complete for every single functional proof predicate for which the axiom of monotonicity **A8** is correct.

4.2 Definition. A model $(K, \prec, \Vdash)$ is **monotone** if in addition to the forcing conditions 1-5 the following forcing condition holds for every node $x \in K$:

6. $x \nVdash \Box_{q_1} A_2(q_2) \& \Box_{q_2} A_3(q_3) \& \ldots \& \Box_{q_n} A_1(q_1)$,

where $A_i(q_i)$ is a modal formula in which the proof variable q_i occurs $(i = 1, \ldots n)$.
(monotonicity)

4.3 Theorem.

For every modal formula A

$$\mathcal{M} \vdash A \quad \Longleftrightarrow \quad A \text{ is valid in all monotone } A\text{-sound models}$$

The proof of this theorem also almost typographically repeats that of Theorem 3.2. One has only to incorporate the monotonicity condition into the definition of a model. As in 3.2 we define the axiom system $\mathcal{M}^-$ as $\mathcal{M}$ without the liberalization rule **R3**, and prove the completeness of $\mathcal{M}^-$ with respect to finite tree functional models with the monotonicity forcing condition 6. This condition also doesn't spoil the algorithm $\mathcal{V}$; we should only insert a monotonicity checker to work on each step of $\mathcal{V}$. Let $\mathcal{V}$ succeed on a frame $(K, \prec)$ with a partial forcing relation $\Vdash_0$ defined on $(K, \prec)$ and SbA, terminates after k steps with a resulting substitution $\tau = \tau_0 \tau_1 \cdots \tau_k$ and a resulting partial forcing relation $\Vdash_k$ satisfying forcing conditions 1-6 for $SbA\tau$, and $SbA\tau$ functional for $(K, \prec \Vdash_k)$.

As before we define the forcing relation $\Vdash$ on $(K, \prec)$ first for atomic and q-atomic formulas Q by

$$x \Vdash Q \qquad \text{iff} \qquad Q\tau \in (SbA)\tau \text{ and } x \Vdash_k Q\tau$$

for every $x \in K$, and then extend it to all modal formulas by forcing conditions 1 and 2.

4.4 Lemma. $(K, \prec, \Vdash)$ is a monotone model.

Proof. In addition of the proof of Lemma 3.9 we have only to demonstrate that the monotonicity forcing condition for $\Vdash$ takes place. Indeed, imagine for some $x \in K$

$$x \Vdash \Box_{q_1} A_2(q_2) \& \Box_{q_2} A_3(q_3) \& \ldots \& \Box_{q_n} A_1(q_1).$$

Define $r_i = q_i\tau$, $B_i(r_i) = (A_i(q_i))\tau$ and note that r_i indeed occurs in $B_i(r_i)$. Then by the definition of $\Vdash$ on q-atomic formulas

$$\Box_{r_1} B_2(r_2), \Box_{r_2} B_3(r_3), \ldots, \Box_{r_n} B_1(r_1) \in SbA\tau$$

and

$$x \Vdash_k \Box_{r_1} B_2(r_2) \& \Box_{r_2} B_3(r_3) \&, \ldots, \& \Box_{r_n} B_1(r_1),$$

which violates the monotonicity condition for $\Vdash_k$.
■

4.5 Corollary. $\mathcal{M}$ is decidable with the computational time of order 2^{cn}, with n the length of the formula, and c fixed.

4.6 Corollary. If $\mathcal{M} \nvdash A$, then there is a substitution τ and an $A\tau$-sound monotone countermodel $\mathcal{K}$ for $A\tau$ such that $(SbA)\tau = Sb(A\tau)$ and $(SbA)\tau$ is functional in $\mathcal{K}$.

4.7 Theorem.

$$\mathcal{M} \vdash A \iff \text{for every Gödel interpretation } * \quad \mathbf{PA} \vdash A^*$$

Proof. Correctness (i.e. the case $\Longrightarrow$) is straightforward by induction on a proof of A in $\mathcal{M}$ like in Theorem 3.14. The monotonicity axiom holds because of our convention on the monotonicity of the Gödel numbering (above).

We proceed now with the proof of the completeness (i.e. the direction $\Longleftarrow$), which also includes the Solovay construction for $\Box$ steps of the induction, but is totally different from that for $\mathcal{F}$ in "labeled modalities" part. Let $\mathcal{M} \nvdash A$. By 4.6 take a substitution τ and a finite $A\tau$-sound monotone countermodel $\mathcal{K}$ for $A\tau$ such that $(SbA)\tau = Sb(A\tau)$ is functional in $\mathcal{K}$. As before we assume that $K = \{1, \ldots, n\}$ and 1 is the root node and define a new model $\mathcal{K}'$ by adding a node 0 to K, putting $0 \prec i$ $(1 \leq i \leq n)$, and defining $0 \Vdash B$ iff $1 \Vdash B$ for every formula $B \in SbA\tau$. It is clear that as in 3.14 the relation $\Vdash$ satisfies forcing conditions 1-6 on $SbA\tau$ at the node 0 as well.

Now for the model $\mathcal{K}'$ and for the usual Gödel proof predicate $\mathit{Proof}(x,y)$ we define a Solovay function $h(t)$, "$l = j$" as a natural arithmetical formula for "j is a limit of $h(t)$", and put

$$\varphi(S_i) := [\bigvee_{j \Vdash S_i} \text{"}l = j\text{"}] \wedge i = i.$$

We begin a process of defining the interpretation $\sharp = (\mathit{Proof}, \varphi)$ (i.e. defining the interpretation of proof variables). Let Var denote the set of all proof variables occurring in $SbA\tau$. A binary relation *less* on Var is defined as follows: q *less* p if there are $\Box_p A_1(q_1), \Box_{q_1} A_2(q_2) \ldots \Box_{q_n} A_{n+1}(q) \in SbA\tau$ such that each $A_i(q_i)$ contains an occurrence of q_i (q occurs in $A_{n+1}(q)$) and

$$\Box_p A_1(q_1) \& \Box_{q_1} A_2(q_2) \& \ldots \& \Box_{q_n} A_{n+1}(q)$$

is valid in $\mathcal{K}'$. It is easy to see that *less* is transitive and irreflexive, and that for every sequence

$$q_1 \ \mathit{less} \ q_2 \ \mathit{less} \ldots \mathit{less} \ q_k$$

k is less or equal to the cardinality of Var. For every $p \in \mathit{Var}$ we define a natural number $rk(p)$:

1. If there is no such $\Box_p F \in SbA\tau$ that $\Box_p F$ is valid in $\mathcal{K}'$, then $rk(p) = 0$.
2. If there is $\Box_p F \in SbA\tau$ such that $\Box_p F$ is valid in $\mathcal{K}'$, then $rk(p) = \max\{k \mid q_1 \ \mathit{less} \ q_2 \ \mathit{less} \ldots \ \mathit{less} \ q_k = p$ for some $q_1, q_2, \ldots, q_k \in \mathit{Var}\}$.

For a formula $F \in SbA\tau$ put $rk(F) = \max\{rk(p) \mid p$ occurs in $F\}$. The following lemma obviously holds

4.8 Lemma. For $\Box_p F \in SbA\tau$ and $j \in K'$,

$$\text{if } j \Vdash \Box_p F, \text{ then } rk(F) < rk(p).$$

Now by induction on k we

1. define an interpretation $\sharp$ for all $F \in SbA\tau$ such that $rk(F) \leq k$;
2. prove that for all $F \in SbA\tau$ such that $rk(F) \leq k$ and for every node $j = 0, 1, \ldots, n$
$$j \Vdash F \implies \mathtt{PA} \vdash \text{"}l = j\text{"} \to F^\sharp,$$
$$j \not\Vdash F \implies \mathtt{PA} \vdash \text{"}l = j\text{"} \to \neg F^\sharp.$$

Basis of the induction, $k = 0$.

1. If suffices to define $\sharp$ for all proof variables p_i such that $rk(p_i) = 0$. Let $rk(p_i) = 0$. It means that $\Box_{p_i} F$ does not hold in $\mathcal{K}'$ for any $\Box_{p_i} F \in SbA\tau$. Put

$$p_i{}^\sharp := \text{ the Gödel number of the shortest proof of } i = i.$$

Now the interpretation $\sharp$ is defined for all formulas F such that $rk(F)=0$.

2. We proceed with the Solovay induction on formulas F such that $rk(F)=0$ as in Lemma 2.15, taking a special care of the case $j=0$. The set of such formulas is closed under subformulas and the induction goes smoothly. The case of sentence variables and $\bot$ is trivial. If $F\equiv\Box_{p_i}G$, then $rk(p_i)=0$, and the situation $j\Vdash\Box_{p_i}G$ is impossible. Then $j\Vdash\neg\Box_{p_i}G$ for all nodes $j\in K'$. But

$$(\Box_{p_i}G)^\sharp = \mathit{Proof}({p_i}^\sharp,\ulcorner G^\sharp\urcorner),$$

which is false as no $G^\sharp$ coincides with $i=i$,

$$\mathbf{PA}\vdash\neg\mathit{Proof}({p_i}^\sharp,\ulcorner G^\sharp\urcorner),$$

and $\mathbf{PA}\vdash$ "$l=j$" $\to\neg F^\sharp$. The induction steps on F go exactly as in 2.15.

Let us now make a step from k to $k+1$.

1. If $rk(p_i)=k+1$, then $j\Vdash\Box_{p_i}F$ for some F, and by Lemma 4.8 $rk(F)<rk(p_i)$, i.e. $rk(F)\leq k$. By the induction hypothesis $F^\sharp$ has already been defined and (2) holds for F. The q-reflexivity and stability of $\mathcal{K}'$ imply $j\Vdash F$ for all $j\in K'$, and by the induction hypothesis $\mathbf{PA}\vdash$ "$l=j$" $\to F^\sharp$ for all $j\in K'$. Thus

$$\mathbf{PA}\vdash\ [\bigvee_{j=0}^{j=n}\text{"}l=j\text{"}]\to F^\sharp,$$

$$\text{by Lemma 2.14 (1)}\quad\mathbf{PA}\vdash\bigvee_{j=0}^{j=n}\text{"}l=j\text{"},$$

thus $\mathbf{PA}\vdash\ F^\sharp$. Now we put

${p_i}^\sharp=$ the Gödel number of the shortest proof of $F^\sharp$ in $\mathbf{PA}$.

2. Again, the Solovay proof works with the new case in its basis: $F\equiv\Box_{p_i}G$ with $rk(p_i)=k{+}1$. If $j\Vdash F$, then again $rk(G)<rk(p_i)$, $rk(G)\leq k$ and by (1), $\mathit{Proof}({p_i}^\sharp,\ulcorner G^\sharp\urcorner)$ holds, $\mathbf{PA}\vdash\mathit{Proof}({p_i}^\sharp,\ulcorner G^\sharp\urcorner)$, and thus $\mathbf{PA}\vdash$ "$l=j$" $\to F^\sharp$.

If $j\not\Vdash F$, then $j\Vdash\Box_{p_i}H$ for some $H\not\equiv G$ with $rk(H)\leq k$ (because $rk(p_i)>0$). By the injectivity of $\sharp$, $F^\sharp\not\equiv H^\sharp$, by the definition of ${p_i}^\sharp$, $\mathit{Proof}({p_i}^\sharp,\ulcorner H^\sharp\urcorner)$ is true and, by the functionality of $\mathit{Proof}(x,y)$, $\mathit{Proof}({p_i}^\sharp,\ulcorner G^\sharp\urcorner)$ is false. Then we have $\mathbf{PA}\vdash\neg\mathit{Proof}({p_i}^\sharp,\ulcorner G^\sharp\urcorner)$ and $\mathbf{PA}\vdash$ "$l=j$" $\to\neg F^\sharp$. Finally $\sharp$ is defined for all $F\in SbA\tau$ and the following lemma is proved

4.9 Lemma. For all $F\in SbA\tau$ and for every node j of $\mathcal{K}'$

$$j\Vdash F\quad\Longrightarrow\quad\mathbf{PA}\vdash\text{"}l=j\text{"}\to F^\sharp,$$

$$j\not\Vdash F\quad\Longrightarrow\quad\mathbf{PA}\vdash\text{"}l=j\text{"}\to\neg F^\sharp.$$

Finally, as in 3.14 we conclude that $\mathtt{PA} \nvdash (A\tau)^\sharp$, define $* := \tau\sharp$ and get the desired $\mathtt{PA} \nvdash A^*$.

$\blacksquare$

4.10 Note. Again the set of $\mathcal{L}^+$ formulas that are true under every Gödel interpretation can be axiomatized by the system $\mathcal{M}'$ whose axioms are the set of theorems of $\mathcal{M}$ together with the scheme $(\Box A \to A)$ and the only rule is *modus ponens.*

5 Acknowledgements

Thanks to Professor G.Jäger who initiated studies in Logic of Proofs, and to T.Strassen for many valuable suggestions. Thanks also to Professor M.Habib, Professor A.Nerode, and Professor A.Preller for helpful discussions and for providing excellent research opportunities in 1992-93.

References

[1] S. Artëmov and T. Straßen, "The Basic Logic of Proofs," in *Computer Science Logic* (E. Bögrer, G. Jäger, H. Kleine Büning, and M. Richter, eds.), Lecture Notes in Computer Science, vol. 702 pp. 14–28, 6th Workshop, CSL'92, San Miniato, Italy, October 1992, Springer-Verlag, 1993.

[2] S. Artëmov and T. Straßen, "Functionality in The Basic Logic of Proofs," Tech. Rep. IAM 92-004, Department for Computer Science, University of Berne, Switzerland, January 1993.

[3] R. M. Solovay, "Provability interpretations of modal logic," *Israel Journal of Mathematics*, vol. 25, pp. 287–304, 1976.

[4] S. Artëmov, "Logic of Proofs," Rapport de Recherche R.R LIRMM No.92-078, CNRS–Laboratoire d'Informatique, de Robotique et de Microelectronique de Montpellier, December 1992.

[5] S. Artëmov, "Logic of Functional Proofs," Tech. Rep. 93-47, Mathematical Sciences Institute, Cornell University, July, 1993.

[6] J. Lassez, M. Maher, and K. Marriott, "Unification revisited," in *Foundations of Deductive Databases and Logic Programming* (J. Minker, ed.), ch. 15, pp. 587–625, Morgan Kaufmann Publishers, Inc., 1987.

[7] D. Guaspary and R. M. Solovay, "Rosser sentences," *Annals of Mathematical Logic*, vol. 16, pp. 81–99, 1979.

[8] G. Boolos, *The unprovability of consistency: an essay in modal logic.* Cambridge: Cambridge University Press, 1979.

[9] C. Smoryński, *Self-reference and modal logic.* New York, Berlin, Heidelberg, Tokyo: Springer Verlag, 1985.

[10] G. E.Hughes and M. J.Cresswell, *A Companion to Modal Logic.* London: Methuen, 1984.

Submitted: December 31, 1992,

Revised: August 10, 1993