Relative Interpretations in Constructive Arithmetic

Albert Visser
Department of Philosophy, Utrecht University
Heidelberglaan 8, 3584 CS Utrecht, The Netherlands
email: Albert.Visser@phil.uu.nl

September 17, 2005

Abstract

In this paper, we show that the predicate logics of consistent extensions of Heyting's Arithmetic plus Church's Thesis with uniqueness condition are complete Π^0_2. Similarly, we show that the predicate logic of HA^*, i.e. Heyting's Arithmetic plus the Completeness Principle (for HA^*) is complete Π^0_2. These results extend the known results due to Valery Plisko. To prove the results we adapt Plisko's method to use Tennenbaum's Theorem to prove 'categoricity of interpretations' under certain assumptions.

Key words: Relative interpretations, predicate logics of arithmetical theories, constructive logic

MSC2000 codes: 03B20, 03B25, 03F25, 03F30, 03F45

Contents

1 Introduction

In this paper, we extend Plisko's negative results concerning predicate logics of constructive theories. We show, for two classes of theories, that the set of schematically valid predicate logical sentences of those theories is complete Π^0_2. The first class consists of theories which prove a version of Church's Thesis.[1] The second class contains sufficiently sound theories T that prove their own Completeness Principle: if A, then A is provable in T.

The methodology we employ is to prove a version of Tennenbaum's Theorem for interpretations, showing that, over the given theory T, a suitable weak arithmetic F is categorical for interpretations. This means that T proves, for any two interpretations, that *if* they both interpret F *then* they are isomorphic. This method was developed by Valery Plisko in a series of papers, to wit [Pli73, Pli77, Pli78, Pli83, Pli90, Pli91, Pli93, Pli02].

There is an illuminating way of looking at these results. We can view the constructive connectives as an *extension* of the classical ones. Let's put ourselves at the standpoint that the constructive connectives are basic and that the classical ones are defined via the double negation translation. From this point of view, classical conjunction, implication and universal quantification coincide with their constructive counterparts, but classical disjunction is $\neg\,(\neg \wedge \neg)$ and classical universal quantification is $\neg\exists\neg$, Thus, constructive disjunction and existential quantification appear as *new connectives* additional to the classical ones. We view e.g. HA as an extension of PA. Now the demand that relative interpretations commute with the new connectives appears as an extra constraint. It is not at all surprising that under this further constraint we will find less relative interpretations.

We could also extend our classical language not with constructive disjunction and existential quantification, but with, say, a modal necessity operator and see what happens then. In fact, this idea was extensively studied for the interpretation of the necessity operator as formal provability. It turned out that we meet analogous phenomena. E.g., following a similar path, we can show that the predicate provability logic of PA is complete Π^0_2.[2]

[1] Our result for the first class improves Plisko's result by eliminating the use of Markov's Principle.

[2] For an exposition, see [Boo93] or [AB04]. The original papers are [Art80, Mon87, Var86, BM87]. For a further development, see [Vis03].

Remark 1.1 There is a clear connection between interpretations and models. A relative interpretation of V in U can be viewed as a uniform construction of V-models as inner models of U-models. Thus, it is not surprising that an analogous categoricity theorem can be proved for models. It is consistent with reasonable constructive metatheories like CZF that Heyting Arithmetic (HA) is categorical.[3] What is more: in a constructive metatheory, like CZF, rich enough to formalize the notion of satifaction, enriched with a suitable version of Church's Thesis and Markov's Principle, we can prove that (a sufficiently rich subsystem of) Heyting's Arithmetic (HA) is categorical. See e.g. [McC91]. It seems to me that Markov's Principle is superfluous: mimicking the methods of this paper one should be able to show, in a sufficiently strong metatheory which includes Church's thesis, that $i\mathsf{S}^1_2$, i.e. the constructive version of Buss' theory S^1_2, is categorical. I will not explore this line further in this paper. ❑

Acknowledgements

I thank Peter Aczel, Lev Beklemishev, Emil Jerabek, Helmut Kohlenbach, Joan Moschovakis and Jaap van Oosten for illuminating discussions.

2 Relative Interpretations

In my paper [Vis04], I developed a framework for translations and interpretations. In the present paper, we need only a small part of that framework. We introduce the notions we need.

2.1 Constructive Predicate Logic

We study the predicate logics of certain arithmetical theories. For that reason it is important to be clear what *predicate logic* is supposed to be. We choose to study constructive predicate logic in relational languages of finite signature with identity. Informally, we will also work with languages with terms. These languages can be translated using a standard algorithm to corresponding relational languages. This algorithm, fortunately, is also correct for constructive logic.

A signature Σ is a triple $\langle \mathsf{pred}, \mathsf{ar}, E\rangle$, where pred is a finite set of predicate symbols, where $\mathsf{ar} : \mathsf{pred} \to \omega$ is the arity function. The special predicate E is binary. It represents the identity relation. We will often write '=' for: E.

We assume that we are given a fixed ω-ordered sequence of variables $v_0, v_1, \ldots$. We will use $x, y, x_0, \ldots$ as metavariables ranging over variables. (We will follow the usual convention that, if e.g. "v_3", "x" and "y" are used in one formula, they are supposed to be distinct.) Formulas and sentences based on Σ and $v_0, v_1, \ldots$ are defined in the usual way.

A theory of signature Σ will be given by its axioms, i.e. a set of sentences of the signature. Derivability from the theory employs the axioms and the rules of

[3] *Models* here are supposed to be models in the classical sense.

constructive predicate logic including the identity axioms and rules for E. We will assume that the axiom set of our theories is appropriately simple, say p-time decidable. The minimal theory of a given signature Σ is the (constructive) predicate logic Pred_Σ.

2.2 Relative Translations

Let Σ and Θ be signatures. A *relative translation* $\tau : \Sigma \to \Theta$ is given by a pair $\langle \delta, \Phi \rangle$. Here δ is a Θ-formula representing the *domain* of the translation. We demand that δ contains at most v_0 free. The mapping Φ associates to each relation symbol R of Σ with arity n an Θ-formula $\Phi(R)$ with variables among $v_0, \ldots, v_{n-1}$. We translate Σ-formulas to Θ-formulas as follows:

- $(R(y_0, \cdots, y_{n-1}))^\tau := \Phi(R)(y_0, \cdots, y_{n-1})$;[4]
- $(\cdot)^\tau$ commutes with the propositional connectives;
- $(\forall y\, A)^\tau := \forall y\, (\delta(y) \to A^\tau)$;
- $(\exists y\, A)^\tau := \exists y\, (\delta(y) \wedge A^\tau)$.

We call τ *unrelativized* if δ_τ is given by $v_0 = v_0$. Suppose τ is $\langle \delta, \Phi \rangle$. Here are some convenient notations.

- We write δ_τ for δ and F_τ for F.
- We write R_τ for $\Phi_\tau(R)$.
- We write $\vec{x} \in \delta$ for: $\delta(x_0) \wedge \ldots \wedge \delta(x_{n-1})$.
- We write $\forall \vec{x} {\in} \delta\, A$ for: $\forall x_0 \ldots \forall x_{n-1}\, (\vec{x} \in \delta \to A)$.
- We write $\exists \vec{x} {\in} \delta\, A$ for: $\exists x_0 \ldots \exists x_{n-1}\, (\vec{x} \in \delta \wedge A)$.

We can compose relative translations as follows:

- $\delta_{\tau\nu} := (\delta_\nu \wedge (\delta_\tau)^\nu)$,
- $R_{\tau\nu} = (R_\tau)^\nu$.

We write $\nu \circ \tau := \tau\nu$. Note that $(A^\tau)^\nu$ is provably equivalent in predicate logic to $A^{\tau\nu}$. The identity translation $\mathsf{id} := \mathsf{id}_\Theta$ is defined by:

- $\delta_{\mathsf{id}} := (v_0 E v_0)$,
- $R_{\mathsf{id}} := R(v_0, \ldots, v_{n-1})$.

[4]Here $\Phi(R)(y_0, \cdots, y_{n-1})$ is our sloppy notation the result of substituting the y_i for the v_i in $\Phi(R)$. We assume that some mechanism for α-conversion is built into our definition of substitution to avoid variable-clashes.

2.3 Relative Interpretations

A translation τ supports a *relative interpretation* of a theory U in a theory V, if, for all U-sentences A, $U \vdash A \Rightarrow V \vdash A^\tau$. (Note that this automatically takes care of the theory of identity. Moreover, it follows that $V \vdash \exists v_0\ \delta_\tau$.) We will write $K = \langle U, \tau, V\rangle$ for the interpretation supported by τ. We write $K : U \to V$ for: K is an interpretation of the form $\langle U, \tau, V\rangle$.

When no confusion is possible, we will often write 'R_K' for 'R_{τ_K}', 'A^K' for A^{τ_K}, etc.

Suppose that T and T' have signature Σ and $T \subseteq T'$. Suppose further that $K : U \to V$, $M : V \to W$. We define:

- $\mathcal{E}_{T,T'} : T \to T'$ is $\langle T, \mathsf{id}_\Sigma, T'\rangle$. $\mathsf{id}_T := \mathcal{E}_{T,T}$.
- $M \circ K : U \to W$ is $\langle U, \tau_M \circ \tau_K, W\rangle$.

We identify two interpretations $K, K' : U \to V$ if:

- $V \vdash \delta_K \leftrightarrow \delta_{K'}$,
- $V, \vec{v} : \delta \vdash P_K \leftrightarrow P_M$, where $\mathsf{ar}(P) = n$ and $\vec{v} = v_0, \ldots, v_{n-1}$.

One can show that modulo this identification, the above operations give rise to a category of interpretations that we call $i\mathsf{INT}$.

We will be interested in interpretations of Pred_Σ in a theory U. It should be noted that it is not automatic that a translation between the right signatures supports such an interpretation. What is needed is precisely that U verifies that δ_τ is inhabited and that E_τ is a congruence (on δ_τ) w.r.t. the translated predicates P_τ. All of this can be expressed in a single sentence.

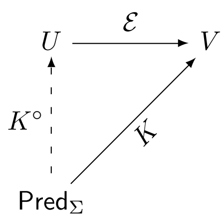
In classical logic, this point is of minor importance. Let's write '$c\mathsf{Pred}$' for classical predicate logic. Consider any translation τ from Σ to Θ. We can effectively find a new translation τ° such that τ° supports an interpretation of $(c)\mathsf{Pred}_\Sigma$ in $c\mathsf{Pred}_\Theta$, and such that, whenever τ supports an interpretation of $(c)\mathsf{Pred}_\Sigma$ in U, then the interpretations from $(c)\mathsf{Pred}_\Sigma$ to U, supported by τ and τ° are the same. The idea is simple. Let C be the single sentence expressing that τ supports an interpretation of predicate logic. Now define:

- $\delta_{\tau^\circ}(v_0) :\leftrightarrow ((C \wedge \delta_\tau(v_0)) \vee (\neg C \wedge v_0 = v_0))$,
- $v_0 E_{\tau^\circ} v_1 :\leftrightarrow ((C \wedge v_0 E_\tau v_1) \vee (\neg C \wedge v_0 = v_1))$,
- $P_{\tau^\circ}\vec{v} :\leftrightarrow P_\tau \vec{v}$.

So, in the classical case, we could always work with the modified translations. These would always give us an interpretation of predicate logic. However, the correctness of this construction relies heavily on classical logic. Thus, in the constructive case, we must take care not to confuse translations with interpretations of predicate logic. Note that we have proved the following lemma.

Lemma 2.1 *Suppose U and V are classical theories and U is a subtheory of V. Suppose further that $K : \mathsf{Pred}_\Sigma \to V$. Then we can find a $K^\circ : \mathsf{Pred}_\Sigma \to U$, such that $K = \mathcal{E}_{U,V} \circ K^\circ$.*

The following picture shows the situation of Lemma 2.1



2.4 Definable Maps between Interpretations

We extend $i\mathsf{INT}$ with extra structure. In this enriched category, $i\mathsf{INT}^{\mathsf{morph}}$, the arrows between two objects play themselves the role of objects in a category as follows. Consider $K, M : U \to V$. An arrow $F : K \Rightarrow M$ is a V-definable, V-provable morphism from K to M considered as 'parametrized internal models'. Specifically, this means that a morphism from K to M is given as a triple $\langle K, F, M\rangle$, where F is a formula with the following properties.

- The free variables of F are among v_0, v_1. We write $F(x, y)$ or xFy, for: $F[v_0 := x, v_1 := y]$.
- $V \vdash xFy \to (x \in \delta_K \wedge y \in \delta_M)$.
- $V \vdash (x \in \delta_K \wedge y \in \delta_M \wedge xE_Kx'Fy'E_My) \to xFy$.
- $V \vdash \forall x{\in}\delta_K\, \exists y{\in}\delta_M\; xFy$.
- $V \vdash (xFy \wedge xFy') \to yE_My'$.
- $V \vdash \vec{x}F\vec{y} \to (P_K\vec{x} \to P_M\vec{y})$.[5]
  Here '$\vec{x}F\vec{y}$' abbreviates $x_0Fy_0 \wedge \ldots \wedge x_{n-1}Fy_{n-1}$, for appropriate n.

We will call the arrows between interpretations: *i-maps*. We consider $F, G : K \Rightarrow M$ as *equal* when they are V-provably the same. The identity $\mathsf{ID}_K : K \Rightarrow K$ is given by: $v_0(\mathsf{ID}_K)v_1 :\leftrightarrow v_0, v_1 \in \delta_K \wedge v_0E_Kv_1$. If $A : K \Rightarrow L$ and $B : L \Rightarrow M$, then $B \cdot A : K \Rightarrow M$ is the obvious 'vertical' composition of B and A.

An isomorphism of interpretations is easily seen to be a morphism with the following extra properties.

- $V \vdash \forall y{\in}\delta_M\, \exists x{\in}\delta_K\; xFy$,
- $V \vdash (xFy \wedge x'Fy) \to xE_Kx'$,

[5] Note that if P represents a function in U, then, by elementary reasoning, we have: $V \vdash \vec{x}F\vec{y} \to (P_K\vec{x} \leftrightarrow P_M\vec{y})$.

- $V \vdash \vec{x}F\vec{y} \to (P_M\vec{y} \to P_K\vec{x})$.

In some cases we can define meaningful *horizontal* compositions between i-maps and interpretations. However, this idea is beyond the scope of the present paper. The reader is referred to [Vis04].

We may obtain a new category $i\mathsf{hINT}$ by dividing out i-isomorphisms between morphisms.

2.5 Predicate Logics of Theories

Consider any theory U. Let Σ be a signature and let Pred_Θ be predicate logic in the signature Σ. We define:

- $\mathsf{PRED}_\Sigma(U) := \{A{\in}\mathsf{sent}_\Sigma \mid \forall K\ ((K : \mathsf{Pred}_\Sigma \to U) \Rightarrow U \vdash A^K)\}$.

We will speak of *the predicate logic of* U. Note that, strictly speaking, the predicate logic of U is a function from signatures to sets of sentences. We will measure *the complexity of the predicate logic of* U by considering a signature for which this complexity is maximal. (There will always be such a signature.)

The unrelativized predicate logic of U is what we obtain if we restrict ourselves in the definition of PRED to unrelativized interpretations.

There is an obvious generalization of our definition. Let U and V be theories. We define:

- $\mathsf{VAL}_V(U) := \{A{\in}\mathsf{sent}_V \mid \forall K\ ((K : V \to U) \Rightarrow U \vdash A^K)\}$.

So, $\mathsf{VAL}_V(U)$ is the set of all U-admissible consequences of V.

We have the following theorem.

Theorem 2.2 *Suppose V is a classical theory and U is a subtheory of V. Then,* $\mathsf{PRED}_\Sigma(U) \subseteq \mathsf{PRED}_\Sigma(V)$.

Proof

Suppose A is in $\mathsf{PRED}_\Sigma(U)$ and $K : \mathsf{Pred}_\Sigma \to V$. By Lemma 2.1, we can find a $K^\circ : \mathsf{Pred}_\Sigma \to U$, such that, for any sentence B of signature Σ, we have $U \vdash B^{K^\circ} \Rightarrow V \vdash B^K$. Clearly, $U \vdash A^{K^\circ}$. Hence, $V \vdash A^K$. ❑

It would be interesting to know a bit more of when we can have monotonicity in the constructive case.

2.6 Preliminaries on Arithmetical Theories: Part 1

In this subsection, we will introduce the preliminaries on arithmetical theories needed for the paper as a whole. For Section 7, we will need some further notions and notations. We postpone their introduction to Subsection 7.1.

Suppose that U is an extension of HA in the language of HA. Suppose also that $K : V \to U$, where V is a subtheory of U which extends $i\mathsf{Q}$, the constructive version of Robinson's Arithmetic. We define in U (modulo arithmetization):

- xF_Kx' iff there is a sequence σ with $\mathsf{length}(\sigma) = x + 1$, such that $\mathsf{Z}_K(\sigma_0)$, and for all $i < x$ we have $\sigma_i \mathsf{S}_K \sigma_{i+1}$ and $\sigma_x = x'$.

We have the following theorem.

Theorem 2.3 *Suppose V and U are arithmetical theories. Suppose further that U satisfies full induction and that $i\mathsf{Q} \subseteq V \subseteq U$. Let $K : V \to U$. Then, F_K is an initial i-embedding from $\mathcal{E}_{V,U}$ to K, i.e. F_K is an injective i-morphism which embeds $\mathcal{E}_{V,U}$ into an initial part of K. Moreover, modulo provable sameness, F_K is the unique i-morphism from $\mathcal{E}_{V,U}$ to K.*

The proof is a formalization of the standard Dedekind argument that the standard numbers are initial in any non-standard model.[6]

We introduce the following useful notation:

(†) $(A\breve{x})^K :\leftrightarrow \exists x' \, (xF_Kx' \wedge (Ax')^K)$.

Note that $U \vdash xF_Ky \leftrightarrow (\breve{x} = y)^K$. We may view our notation as follows. First, we extend the predicate logical language with a new variable $\breve{x}$ for each variable x. We stipulate that the $\breve{x}$ may only occur as free variables. Now we extend the clauses for interpretations, by (†), stipulating that this clause is always executed *before* the clause for the main connective.

Note that $(B\breve{x} \wedge C\breve{x})^K$ will not be syntactically equal of to $(B\breve{x})^K \wedge (C\breve{x})^K$. However, using the functionality of F_K, we can get the commutation clauses for $(\cdot)^K$, modulo provable equivalence. E.g.

- $U, y \in \delta_K \vdash (A\breve{x}y \wedge B\breve{x}y)^K \leftrightarrow ((A\breve{x}y)^K \wedge (B\breve{x}y)^K)$,
- $U, y \in \delta_K \vdash (\neg\, A\breve{x}y)^K \leftrightarrow \neg\, (A\breve{x}y)^K$.
- $U, y \in \delta_K \vdash (\forall z \, A\breve{x}yz)^K \leftrightarrow \forall z{\in}\delta_K \, (A\breve{x}yz)^K$.

From the fact that F is an initial embedding, we have that, for Σ^0_1-formulas $S\vec{x}$,

- $U \vdash S\vec{x} \to (S\breve{\vec{x}})^K$.

[6]Suppose U extends $i\mathsf{Q}$. Then, the statement that, for every $K : i\mathsf{Q} \to U$, there is a unique morphism from $\mathcal{E}_{i\mathsf{Q},U}$ to K, is equivalent to induction. I.o.w., $\mathcal{E}_{i\mathsf{Q},U}$ is initial in the i-category of interpretations from $i\mathsf{Q}$ to U iff U satisfies full induction. See [Vis04].

3 The Classical Case

Our focus in this paper is on predicate logics of constructive arithmetical theories. The aim of this section is just to briefly survey what is known in the classical case. We will just treat arithmetical theories. For discussion of the general case, see [Yav97]. We have:

Theorem 3.1 *Consider any finitely axiomatized arithmetical theory U. Then, we can find a recursively axiomatizable extension W of U such that the predicate logic of W is complete Π^0_2.*

This is in essence Theorem (C.7) on p127 of [Vis05]. In case we consider essentially reflexive theories, the situation is different. The following theorem is a special case of the results of Appendix A of [Vis99].

Theorem 3.2 *Let U be an extension of PA in the language of PA. We have: $\mathsf{PRED}_\Sigma(U) \vdash A$ iff $U \vdash \Box_{c\mathsf{Pred}_\Sigma} A$.*

So, the predicate logic of U corresponds precisely with U's idea of what predicate logic is.

Proof

By formalizing the model existence lemma we can find an interpretation $K : B \to (U + \mathsf{con}(B))$. Here consistency is *consistency in classical logic.* By Lemma 2.1, we can find an interpretation $K^\circ : c\mathsf{Pred}_\Sigma \to U$ such that

$$\mathcal{E}_{U,U+\mathsf{con}(B)} \circ K^\circ = K \circ \mathcal{E}_{c\mathsf{Pred}_\Sigma,B}.$$

It follows, taking $B := \neg A$, that, if $U \vdash A^{K^\circ}$, then $U \vdash \Box_{c\mathsf{Pred}_\Sigma} A$.

Conversely, suppose $U \vdash \Box_{c\mathsf{Pred}_\Sigma} A$. Let K be any interpretation of $c\mathsf{Pred}_\Sigma$ in U. We only need finitely many axioms to verify that K is such an interpretation. Let these axioms be below n. We find $U \vdash \Box_{U,n} A^K$. By the essential reflexivity of U, we get $U \vdash A^K$. ❑

We have the following corollary.

Corollary 3.3 *Let U be an extension of PA in the language of PA. We have that $\mathsf{PRED}_\Sigma(U)$ is recursively enumerable. Moreover, the predicate logic of U is ordinary predicate logic iff U is Σ^0_1-sound.*

Proof

We treat the second claim. In case U is Σ^0_1-sound, then $U \vdash \Box_{\mathsf{Pred}_\Sigma} A$ implies $\mathsf{Pred}_\Sigma \vdash A$. Conversely, suppose that $U \vdash S$, where S is a false Σ^0_1-sentence. Let Θ be the signature of arithmetic. Then, we have, for any interpretation K of Pred_Θ in U, $U \vdash (\mathsf{Q} \to S)^K$. So, $\mathsf{PRED}_\Theta(U) \vdash \mathsf{Q} \to S$. ❑

Open Question 3.4 Let U be an extension of PA in the language of PA. It is a bit unsatisfactory that we do not have an axiomatization of $\mathsf{PRED}_\Sigma(U)$. I conjecture that this is $c\mathsf{Pred}_\Sigma$ plus all axioms of the form $(\mathsf{Q} \to S)^K$, where $U \vdash S$ and where $K : c\mathsf{Pred}_\Theta \to c\mathsf{Pred}_\Sigma$. Here Θ is the signature of arithmetic. ❑

4 Categoricity

In this section, we study categoricity for interpretations. We fix some finitely axiomatized theory F, which is not too strong. It is pleasant to keep this theory 'fluid', so that we can add assumptions to it, when needed. For the time being F could be any theory between $i\mathsf{Q}$ and $i\mathsf{EA}$, the constructive version of Elementary Arithmetic, which is given by Δ_0-induction and the axiom Exp, which states that exponention is total. We will always assume that F is a subtheory of the theories we are considering.

We define the Categoricity Scheme Cat and the Categoricity Rule CatR, as follows:

CatS $\quad \vdash \mathsf{F}^K \to$ "F_K is an isomorphism".

CatR $\quad \vdash \mathsf{F}^K \Rightarrow \vdash$ "F_K is an isomorphism".

In the formulation of the scheme and the rule, "K" ranges over interpretations of Pred_Θ in U. Here Θ is the signature of arithmetic.[7]

Both CatS and CatR are inconsistent with classical logic, under rather general circumstances. Suppose U is classical and $U \vdash \mathsf{con}(\mathsf{F})$. In this case, there is a $K : \mathsf{F} \to U$ that is *restricted*, i.e. such that there is a predicate $\widetilde{K}$ with $U \vdash A^K \leftrightarrow \widetilde{K}(\ulcorner A \urcorner)$, for all arithmetical sentences A. Using the supposed isomorphism F_K, we may transform $\widetilde{K}$ into an ordinary truthpredicate. Thus, both CatS and CatR are in immediate contradiction with Tarski's theorem of the undefinability of truth.

We will consider the Scheme and the Rule for extensions of HA in the language of HA. Since, $\mathsf{HA} \vdash \mathsf{con}(i\mathsf{EA})$ and $i\mathsf{EA}$ will be the strongest version of F we consider, there are no classical extensions of HA satisfying CatS or CatR.

Open Question 4.1 Are there interesting equivalents or characterizations of CatS and CatR? ❑

We sketch an argument due to Valery Plisko. Suppose that U is a consistent extension of HA in the arithmetical language and that U implies CatS. We show that the predicate logic of U is complete Π^0_2.

To avoid cluttering the presentation, let's first suppose that U is Σ^0_1-sound. We consider relative interpretations of predicate logic for Θ_X, the signature of

[7]Note that my use of 'Scheme' versus 'Rule' is *abus de langage*. In fact, the Scheme is also a rule, since the Scheme is only applicable if $U \vdash C_{\tau_K}$, where C_{τ_K} is the single axiom expressing that τ_K supports an interpretation of predicate logic.

the arithmetical language plus one additional unary predicate X. We define, in this language, a formula $\mathsf{TP}(x, X)$ expressing that 'X is a truth predicate for arithmetical formulas up to complexity x' (in the sense that it commutes with the predicate logical connectives in the desired way). Here we take any system of computing complexities such that in HA we have (i) a partial truthpredicate for formulas of any complexity and (ii) a complexity for each formula.[8] Consider any Π^0_2-sentence $A = \forall x\,\exists y\; A_0xy$, where A_0 is Δ^0_0. We define $\widetilde{A}$ as follows:

- $\widetilde{A} :\leftrightarrow (\bigwedge \mathsf{F} \to \exists u\,(\neg\,\mathsf{TP}(u, X) \wedge \forall x{<}u\,\exists y\; A_0xy))$.

We claim that $\widetilde{A}$ is in the predicate logic of U iff A is true.

First, suppose that A is true. Consider any $K : \mathsf{Pred}_{\Theta_X} \to U$. Let K^- be the restriction of K to the signature Θ. Thus, $K^- : \mathsf{Pred}_\Theta \to U$. By CatS, we find that $F := F_{K^-}$ is an i-isomorphism in $U + \mathsf{F}^K$. We have, for any n and any arithmetical sentence C of complexity $< n$ with Gödelnumber c,

$$U + \mathsf{F}^K \vdash (\mathsf{TP}(\underline{\breve{n}}, X) \to (X(\underline{\breve{c}}) \leftrightarrow C))^K \tag{1}$$

(Here, we read the underlining with wide scope.) Let $B := X^K$. It follows that:

$$U + \mathsf{F}^K \vdash (\mathsf{TP}(\underline{\breve{n}}, X))^K \to (\exists x\,(\underline{c}Fx \wedge Bx) \leftrightarrow C) \tag{2}$$

Using the Gödel Fixed Point Lemma, we can construct a formula L with Gödelnumber ℓ such that:

$$\mathsf{HA} \vdash L \leftrightarrow \neg\,\exists x\,(\underline{\ell}Fx \wedge Bx) \tag{3}$$

We take N to be the complexity of C plus 1. We may conclude, combining equation (2) (with L, ℓ, N for C, c, n) with equation (3):

$$U + \mathsf{F}^K \vdash \neg\,(\mathsf{TP}(\underline{\breve{N}}, X))^K \tag{4}$$

We clearly have $\mathsf{F} \vdash \forall x{<}\underline{N}\,\exists y\; A_0xy$. So, $U \vdash (\bigwedge \mathsf{F} \to \forall x{<}\underline{N}\,\exists y\; A_0xy)^K$. Thus, $U \vdash (\widetilde{A})^K$. We find that $\widetilde{A}$ is in the predicate logic of U.

Conversely, suppose $\widetilde{A}$ is in the predicate logic of U. Consider any n. We define e_n as follows. The interpretation e_n is unrelativized. When restricted to Θ it is the identical interpretation. Moreover, e_n sends X to a partial truthpredicate for formulas of complexity n. Specializing the U-validity of $\widetilde{A}$ to the interpetation e_n, it follows that $U \vdash \forall x{<}\underline{n}\,\exists y\; A_0xy$. By Σ^0_1-soundness we are done.

To handle the non-Σ^0_1-sound case, we use the following theorem, which is a minor variation of a theorem due independently to Friedman, Goldfarb and Harrington.

Theorem 4.2 *Let U be a consistent arithmetical theory that extends $i\mathsf{Q}$. For any Σ^0_1-formula Sx, we can effectively find a Σ^0_1-formula Rx, such that $S\underline{n}$ is true iff $R\underline{n}$ is true iff $U \vdash R\underline{n}$.*

[8] See e.g. [Bur00] for good notion of complexity.

See [Vis05] for an exposition. Now take Sx to be $\exists y\ A_0xy$. Let Rx be as promised by the above theorem and take $B := \forall x\ Rx$. Define $A^\circ := \widetilde{B}$. Then we clearly have that A is true if B is true. Moreover, since $U \vdash \forall x{<}\underline{n}\, Rx$ implies that $\forall x{<}\underline{n}\, Rx$ is true, we find that A° is in the predicate logic of U iff B is true. Combining, we have that A° is in the predicate logic of U iff A is true.

Note that the above result holds also when we restrict the class of interpretations to any class that includes the e_n. Thus, the result holds for unrelativized interpretations.

5 Church's Thesis with Uniqueness Condition

In this section, we will show that, over HA, Church's Thesis with uniqueness condition implies the Categoricity Scheme. Let's for the moment assume that F is $i\mathsf{EA}$. Later, we will argue that with some care F can be taken to be $i\mathsf{S}^1_2$. Let T be Kleene's T-predicate. We define:

- $x \cdot y \simeq z :\leftrightarrow \exists u\ (Txyu \wedge Uu = z)$,
- $(x \cdot y)\downarrow\ :\leftrightarrow \exists u\ Txyu$,
- $x \cdot_v y \simeq z :\leftrightarrow \exists u{<}v\ (Txyu \wedge Uu = z)$.

Let U be an extension of $\mathsf{HA} + \mathsf{CT}_0!$. Here:

$\mathsf{CT}_0!$ $\quad \vdash \forall x\, \exists! y\ Axy \to \exists e\, \forall x\ ((e \cdot x)\downarrow \wedge\ Ax(e \cdot x))$

The scheme $\mathsf{CT}_0!$ is essentially weaker than the scheme CT_0. This has been shown by Vladimir Lifschitz in his [Lif79]. See also [Oos90].

Suppose $K : \mathsf{F} \to U$. Let $F = F_K$ be the unique initial embedding. We will show that F is surjective and, hence, an i-isomorphism. Thus K will be i-isomorphic to $\mathcal{E}_{\mathsf{F},U}$.

By our stipulation that F is $i\mathsf{EA}$, we have:

$$(\ddagger)\ \ \mathsf{F} \vdash \forall a\, \forall x\ (\neg\,(x \cdot_a x \simeq 0) \vee (x \cdot_a x \simeq 0)).$$

This follows from the following facts.

a) $i\mathsf{EA}$ proves $\Delta_0(\mathsf{exp})$-induction.

b) $\Delta_0(\mathsf{exp})$-induction proves decidability of $\Delta_0(\mathsf{exp})$-formulas.

c) The T-predicate is $\Delta_0(\mathsf{exp})$.

Reason in U. Consider any $a \in \delta_K$. We have:

$$(\forall x\ (\neg\,(x \cdot_a x \simeq 0) \vee (x \cdot_a x \simeq 0)))^K \qquad (5)$$

So, it follows:

$$\forall x\ (\neg(\breve{x} \cdot_a \breve{x} \simeq 0)^K \vee (\breve{x} \cdot_a \breve{x} \simeq 0)^K) \tag{6}$$

By $\mathsf{CT}_0!$ we can find a recursive index e such that:

$$\forall x \qquad \exists y{\leq}1\ \ e \cdot x \simeq y \tag{7}$$

$$\forall x \qquad e \cdot x \simeq 0 \to \neg(\breve{x} \cdot_a \breve{x} \simeq 0)^K \tag{8}$$

$$\forall x \qquad e \cdot x \simeq 1 \to (\breve{x} \cdot_a \breve{x} \simeq 0)^K \tag{9}$$

Suppose $e \cdot e \simeq 1$. Then, by the initiality of F_K, $(\breve{e} \cdot \breve{e} \simeq 1)^K$. On the other hand, by equation (9), $(\breve{e} \cdot \breve{e} \simeq 0)^K$. A contradiction.

So, $e \cdot e \simeq 0$. Thus, for some v, we have $e \cdot_v e \simeq 0$ and, hence, $(\breve{e} \cdot_{\breve{v}} \breve{e} \simeq 0)^K$. Suppose that $(\breve{v} < a)^K$. It follows that $(\breve{e} \cdot_a \breve{e} \simeq 0)^K$. So, we have a contradiction with equation (8). We may conclude that $(a \leq \breve{v})^K$. Hence, for some z, zFa.

Thus, F is surjective and, hence, an isomorphism.

Note that (‡) is the most heavy assumption we made about F. Since Kleene's T-predicate is Δ^{b}_1 in $i\mathsf{S}^1_2$, the main problem is the first bounded universal quantifier in $(\breve{x} \cdot_a \breve{x} \simeq 0)$. Note, however, that we could replace (‡) by

$$(\pounds) \quad \mathsf{F} \vdash \forall a\, \forall x\ (\neg\,(x \cdot_{|a|} x \simeq 0) \vee (x \cdot_{|a|} x \simeq 0))^K.$$

The principle (£) is verifiable in $i\mathsf{S}^1_2$, since $(x \cdot_{|a|} x \simeq 0)$ will be Δ^{b}_1. Our argument now tells us that $b := |a|$ in K is in the range of F. Say xFb. Now it is easy to see that, there is a c with $2^{x+1}Fc$ and $c >_K a$. Hence, for some y, yFa, and we are done.

Finally, consider any theory W extending $\mathsf{HA}+\mathsf{CT}_0!$. We verify that W satisfies CatS. Consider any $K : \mathsf{Pred}_\Theta \to W$. Let $\tau := \tau_K$ be the underlying translation. Clearly, τ also supports an interpretation $K^{\#} : \mathsf{F} \to (W + \mathsf{F}^\tau)$. Now apply the above result, noting that, qua formula, $F_K = F_\tau = F_{K^{\#}}$.

Remark 5.1 Let I be a formula with just x free, Consider, the following variant of $\mathsf{CT}_0!$.

$$\mathsf{CT}_0!(I)\ \ \vdash (\mathsf{cut}(I) \wedge \forall x{\in}I\ \exists! y{\in}I\ Axy) \to \exists e{\in}I\ \forall x{\in}I\ (((e \cdot x)\downarrow)^I \wedge Ax(e \cdot x))$$

We extend $i\mathsf{EA}$ with all $\mathsf{CT}_0!(I)$. We can adapt the above arguments to show that the resulting theory satisfies full induction, and, thus, coincides with $\mathsf{HA} + \mathsf{CT}_0!$. Thus, in a rather weak sense, Church's Thesis implies induction. ❑

6 Interpretations over HA

In the light of the results of the previous section what can we say about interpretations in HA? Suppose $K : \mathsf{F} \to \mathsf{HA}$. Then, it follows that $\mathcal{E}_{\mathsf{HA},\mathsf{HA}+\mathsf{CT}_0!} \circ K$ is isomorphic to $\mathcal{E}_{\mathsf{F},\mathsf{HA}+\mathsf{CT}_0!}$. Thus, if $\mathsf{HA} \vdash A^K$, then $\mathsf{HA} + \mathsf{CT}_0! \vdash A$. So the arithmetical extensions of F that are interpretable in HA will be bounded by $\mathsf{HA}+\mathsf{CT}_0!$.

It also follows that such a K cannot be restricted, since it is i-isomorphic to $\mathcal{E}$ in $\mathsf{HA}+\mathsf{CT}_0!$.

We can do better if we assume the domain of K to be almost negative. Suppose δ_K in almost negative. We find:

$$\mathsf{HA} \vdash \forall a{\in}\delta_K\ (\forall x\ (\neg\,(x \cdot_{|a|} x \simeq 0) \vee (x \cdot_{|a|} x \simeq 0)))^K \tag{10}$$

Hence:

$$\mathsf{HA} \vdash \forall a{\in}\delta_K\ \forall x\ (\neg\,(\breve{x} \cdot_{|a|} \breve{x} \simeq 0)^K \vee (\breve{x} \cdot_{|a|} \breve{x} \simeq 0)^K) \tag{11}$$

Now we may apply extended Church's Rule to obtain the desired index e of a function that chooses between the disjuncts.[9] We now argue as before to show that F_K is an isomorphism.

Thus, we find that, over HA, the theory HA itself admissibly follows from $i\mathsf{S}^1_2$, provided that we only consider interpretations with almost negative domain formulas. Since, unrelativized interpretation have, a fortiori, an almost negative domain, we have: $\mathsf{VAL}^{\mathsf{unr}}_{i\mathsf{S}^1_2}(\mathsf{HA}) = \mathsf{HA}$.

Open Question 6.1 Can we relatively interpret some strict extension of HA in HA? ❑

7 The Completeness Principle

In this section we study the predicate logic of the theory HA^*. We will introduce the theory in Subsection 7.2.

7.1 Preliminaries on Arithmetical Theories: Part 2

We will employ Guaspari's witness comparison notation. Suppose A is of the form $\exists x\ A_0 x$ and B is of the form $\exists y\ B_0 y$. Suppose further that x is not free in B and y is not free in A.[10] We will write:

- $A \leq B :\leftrightarrow \exists x\ (Ax \wedge \forall y{<}x\ \neg By)$.
- $B < A :\leftrightarrow \exists y\ (By \wedge \forall x{<}y\ \neg Ax)$.
- $\mathsf{E}y :\leftrightarrow \exists z\ y = z$.

[9] We need the extended rule to find e on the assumption that $a \in \delta_K$, for an almost negative δ_K. Note that Troelstra and van Dalen misstate the principle in their [TvD88], p243, forgetting the restriction on the antecedent.

[10] If A and B do not satisfy the variable conditions, we take suitable α-variants that do.

So, we have e.g.: $\mathsf{E}y \leq \exists x\ Ax \leftrightarrow \forall x{<}y\ \neg Ax$. If we have a disjunction $\exists x\ A_0x \vee \exists y\ B_0y$ in a witness comparison, we read this formula as $\exists z\ (A_0z \vee B_0z)$.

We will employ the following arithmetizations.

- sent: the predicate defining the set of (gödelnumbers of) sentences of the arithmetical language.
- $\mathsf{proof}_T(p, x)$: the arithmetization of the proof predicate.
- $\mathsf{prov}_T(x)$: the arithmetization of the provability predicate.
- neg: the arithmetization of the operation that sends a formula to its negation.
- $\underline{n}$: the numeral of n.
- $\ulcorner A \urcorner$: the numeral of the the gödelnumber of A.
- $\mathsf{sub}(x, y, z)$: the relational arithmetization of the function that sends x and y to the gödelnumber z of the formula that is the result of substituting the numeral of x for v_0 in the formula coded by y. So, e.g., the following is a true sentence: $\mathsf{sub}(\underline{n}, \ulcorner v_0 = v_1 \urcorner, \ulcorner \underline{n} = v_1 \urcorner)$;
- $\Box_T A\dot{y}$: $\exists p\, \exists z{<}p\ (\mathsf{sub}(y, \ulcorner Av_0 \urcorner, z) \wedge \mathsf{proof}_T(p, z))$ (and similarly for several variables). We put the quantifier over proofs in front because of the role this quantifier will play in witness comparisons.
- $p : \Box_T A\dot{y}$: $\exists z{<}p\ (\mathsf{sub}(y, \ulcorner Av_0 \urcorner, z) \wedge \mathsf{proof}_T(p, z))$.

We won't consider the dot notation under multiple nestings of the $\Box$, so we need not discuss what happens in that case. To get some feeling for the combination of the dot notation and the breve notation, we look at some examples.

i) $\Box_T(A\check{\dot{x}})^K$: this is $\exists p\, \exists z{<}p\ (\mathsf{sub}(x, \ulcorner \exists v_1(v_0F_Kv_1 \wedge Av_1) \urcorner, z) \wedge \mathsf{proof}_T(p, z))$.

ii) $\Box_T(A\dot{\check{x}})^K$: this makes no sense. To unwind the dot notation, we first have to compute the K-translation. This cannot be done without first removing the dot.

iii) $(\Box_T A\dot{\check{x}})^K$: this is $\exists y\ (xF_Ky \wedge \exists p\, \exists z{<}p\ (\mathsf{sub}(y, \ulcorner Av_0 \urcorner, z) \wedge \mathsf{proof}_T(p, z)))$.

7.2 What is HA^*?

The theory HA^* was introduced in [Vis82]. The theory consists of HA plus *the Completeness Principle for* HA^*. The Completeness Principle for a theory T is given by:

$\mathsf{CP}[T] \quad \vdash A \to \Box_T A$.

We have $\mathsf{HA}^* := \mathsf{HA} + \mathsf{CP}[\mathsf{HA}^*]$.[11] The theory HA^* was used as a technical tool in [dJV96] and in [Vis02].

A principle closely connected to the Completeness Principle is the Strong Löb Principle SLP. This principle is given by:

$\mathsf{SLP}[T] \quad \vdash (\Box_T A \to A) \to A$

As a special case of SLP we have $\vdash \neg\neg\Box_T\bot$. One can show that SLP and CP are interderivable.

We briefly review some of the results of [Vis82] and [dJV96].

- Michael Beeson introduced the notion of fp-realizability in [Bee75]. The simplest variant of this form of realizability is a provability translation. The theory HA^* is to this translation as Troelstra's $\mathsf{HA} + \mathsf{ECT}_0$ is to Kleene's r-realizability. This means that HA^* is the set of sentences such that their provability translations are provable in HA.

- Let $\mathfrak{A}$ be the smallest class closed under atoms and all connectives, where the clause for implication is restricted as follows:

  $$(A \in \Sigma_1 \text{ and } B \in \mathfrak{A}) \Rightarrow (A \to B) \in \mathfrak{A}.$$

  Note that modulo provable equivalence in HA all prenex formulas (of the classical arithmetical hierarchy) are in $\mathfrak{A}$. The theory HA^* is conservative w.r.t. $\mathfrak{A}$ over HA.

- There are infinitely many incomparable T with $T = \mathsf{HA} + \mathsf{CP}[T]$. However if $T = \mathsf{HA} + \mathsf{CP}[T]$ *verifiably in* HA, then $T = \mathsf{HA}^*$.

- Let KLS:=Kreisel-Lacombe-Shoenfield's Theorem on the continuity of the effective operations. We have $\mathsf{HA}^* \vdash \mathsf{KLS} \to \Box_{\mathsf{HA}^*}\bot$. This immediately gives Beeson's result that $\mathsf{HA} \nvdash \mathsf{KLS}$. See [Bee75].

- Every prime RE Heyting algebra $\mathcal{H}$ can be embedded into the Heyting algebra of HA^*. This mapping is primitive recursive and sends every element of the algebra to a Σ-sentence (modulo provability). See [dJV96].

7.3 The Predicate Logic of HA^*

In this subsection, we will show that the predicate logic of HA^* is complete Π^0_2. We will write $\Box^*$ for $\Box_{\mathsf{HA}^*}$, etc. We start with an analogue of Tennenbaum's theorem. We will take $\mathsf{F} := i\mathsf{EA}$.

Theorem 7.1 *We have:*
$\mathsf{HA}^* + \mathsf{F}^K \vdash \forall y{\in}\delta_K\, (\exists z\, (y = \breve{z})^K \vee \Box^*\bot)$.

[11] The natural way to define HA^* is by a fixed point construction as: HA plus the *Completeness Principle* for HA^*. Here it is essential that the construction is verifiable in HA.

Proof

Define the following formula:

- $Buy :\leftrightarrow \exists z\ (\,\mathsf{sub}(y,u,z) \wedge ((\mathsf{E}y \vee \mathsf{prov}^*(\mathsf{neg}(\breve{z}))) \leq \mathsf{prov}^*(\breve{z}))^K)$

Note that Buy is equivalent over $\mathsf{HA}^* + \mathsf{F}^K + y \in \delta_K$ with Cuy, where:

- $Cuy :\leftrightarrow (\exists w\ (\,\mathsf{sub}(\breve{y},\breve{u},w) \wedge (\mathsf{E}y \vee \mathsf{prov}^*(\mathsf{neg}(w))) \leq \mathsf{prov}^*(w)))^K$

By the Gödel Fixed Point Lemma, we can find a formula Rv_0 such that:

$$i\mathsf{Q} \vdash Ry \leftrightarrow B(\ulcorner Rv_0 \urcorner, y) \tag{12}$$

Equation (12) implies:

$$\mathsf{HA}^* + \mathsf{F}^K + y \in \delta_K \vdash Ry \leftrightarrow ((\mathsf{E}y \vee \Box^* \neg\, R\dot{\breve{y}}) \leq \Box^* R\dot{\breve{y}})^K \tag{13}$$

Note the strangeness of the double role that y plays in the right hand side of this equivalence. In the occurrence in '$\mathsf{E}y$' we have y in the role of a domain element of K. In the occurrences with a breve, y is an element of the external numbers mapped into the K-numbers. So, in a sense, y stands for two different K-numbers, to wit y and z with yF_Kz.

Reason in $\mathsf{HA}^* + \mathsf{F}^K + y \in \delta_K$. We clearly have $(\mathsf{E}y)^K$. Hence, since F is $i\mathsf{EA}$, we have Ry or $R^\perp y$, where

- $R^\perp y :\leftrightarrow (\Box^* R\dot{\breve{y}} < (\mathsf{E}y \vee \Box^* \neg\, R\dot{\breve{y}}))^K$

Suppose we have Ry. Then, by the Completeness Principle, we have $\Box^* R\dot{y}$. Let p be a witnessing proof. It follows that $(\breve{p} : \Box^* R\dot{\breve{y}})^K$. By, equation (13), we either have $(y \leq \breve{p})^K$ or $(\Box^* \neg\, R\dot{\breve{y}} \leq \mathsf{E}\breve{p})^K$. In the first case, we find that, for some z, we have $(y = \breve{z})^K$. In the second case, we find that, for some q, we have $(\breve{q} : \Box^* \neg\, R\dot{\breve{y}})^K$ and, hence, $q : \Box^* \neg\, R\dot{y}$. Combining this with $\Box^* R\dot{y}$, we find $\Box^* \bot$.

Now suppose $R^\perp y$. It follows that $\neg\, Ry$. By the Completeness Principle, we find that, for some r, we have $r : \Box^* \neg\, R\dot{y}$. Hence $(\breve{r} : \Box^* \neg R\dot{\breve{y}})^K$. Now, using $R^\perp y$ again, we find $(\Box^* R\dot{\breve{y}} < \mathsf{E}\breve{r})^K$ and, hence, $\Box^* R\dot{y}$. We may conclude $\Box^* \bot$.

Combining our two cases, we see that, for some z, $(y = \breve{z})^K$ or $\Box^* \bot$. ❑

Here is an immediate consequence of Theorem 7.1.

Theorem 7.2 *We have:*
$\mathsf{HA}^* + \vec{x} F_K \vec{y} \vdash \Box^* A\dot{\vec{x}} \leftrightarrow (\Box^* A\dot{\vec{y}})^K$.

Proof

We reason in $\mathsf{HA}^* + \vec{x}F_K\vec{y}$. From left to right is immediate. For the right to left direction, suppose that $(q : \Box^* A\dot{\vec{y}})^K$. By Theorem 7.1, either, for some p, pF_Kq, or $\Box^*\bot$. In the first case we have $p : \Box^* A\dot{\vec{x}}$, in the second case $\Box^* A\dot{\vec{x}}$. So in both cases, we are done. ❑

For a generalization of Theorem 7.2, see Appendix A.

Theorem 7.3 *The predicate logic of* HA^* *is complete* Π^0_2.

Proof

Suppose $A :\leftrightarrow \forall x\, \exists y\ A_0xy$, where A_0 is in Δ^0_1. We will (effectively) construct a formula $\widetilde{A}$, such that $\widetilde{A}$ is in the predicate logic of HA^* iff A is true. The theorem is immediate from this claim. The predicate logical language that we interpret is the relational language of arithmetic enriched by a 0-ary predicate symbol Q. We define:

- $\widetilde{A} :\leftrightarrow ((\mathsf{F} \wedge \exists x{\in}\mathsf{sent}\ (Q \leftrightarrow \mathsf{prov}^*(x))) \to$
$$\exists x, y{\in}\mathsf{sent}\ ((Q \leftrightarrow \mathsf{prov}^*(x)) \wedge (\mathsf{prov}^*(y) \leftrightarrow \Box^*(\mathsf{prov}^*(\dot{x}))) \wedge \forall u{<}y\, \exists v\ A_0uv)).$$

First, suppose $\forall x\, \exists y\ A_0xy$. Consider any K. Let B be the arithmetical sentence Q_K. Let C be the consequent of the implication $\widetilde{A}$. Consider the following theory:

$$T := \mathsf{HA}^* + \mathsf{F}^K + (\exists x{\in}\mathsf{sent}\ (Q \leftrightarrow \mathsf{prov}^*(x)))^K.$$

Note that, for any n, $T \vdash \forall u{<}\underline{n}\, \exists v\ A_0uv$ and also $T \vdash (\forall u{<}\underline{n}\, \exists v\ A_0uv)^K$.

Reason in T. We have to prove C^K. Suppose x witnesses

$$(\exists x{\in}\mathsf{sent}\ (Q \leftrightarrow \mathsf{prov}^*(x)))^K.$$

We have either zF_Kx, for some z, or $\Box^*\bot$. We first treat the second case. Suppose we have $\Box^*\bot$. Let $\underline{n} := \ulcorner\top\urcorner$. Now it is easy to see that we can choose x and y to be equal to $\breve{\underline{n}}$.

We turn to the first case. We find that $z \in \mathsf{sent}$ and, applying Theorem 7.2, $B \leftrightarrow \mathsf{prov}^*(z)$. So, we have $\Box^*B \leftrightarrow \Box(\mathsf{prov}^*(\dot{z}))$, by the Completeness Principle. Thus, taking $\underline{n} := \ulcorner B\urcorner$, we find:

$$z, \underline{n} \in \mathsf{sent} \wedge (B \leftrightarrow \mathsf{prov}^*(z)) \wedge (\mathsf{prov}^*(\underline{n}) \leftrightarrow \Box^*(\mathsf{prov}^*(\dot{z}))) \wedge \forall u{<}\underline{n}\, \exists v\ A_0uv. \quad (14)$$

Using Theorem 7.2 again, from this we immediately obtain C^K.

For the reverse direction, let a_n be unrelativized and let a_n be the identity on all arithmetical predicates and let $Q_{\mathsf{a}_n} := \Box^{*\,n+1}\bot$. Suppose $\widetilde{A}$ is in the predicate logic of HA^*. It follows, by instantiating with a_n, that, for any n,

$$\mathsf{HA}^* \vdash \exists y{\in}\mathsf{sent}\ ((\mathsf{prov}^*(y) \leftrightarrow \Box^{*\,n+2}\bot) \wedge \forall u{<}y\, \exists v\ A_0uv). \quad (15)$$

Since HA^* has the existence property, for some $k = k_n$,

$$\mathsf{HA}^* \vdash \underline{k}{\in}\mathsf{sent} \wedge (\mathsf{prov}^*(\underline{k}) \leftrightarrow \Box^{*\,n+2}\bot) \wedge \forall u{<}\underline{k}\,\exists v\; A_0uv. \tag{16}$$

So, by the Σ^0_1-soundness of HA^*, we find $\forall u{<}\ k\,\exists v\; A_0uv$. Suppose $k_n = k_m$, for $m > n$. Then,

$$\mathsf{HA}^* \vdash \Box^{*\,n+2}\bot \leftrightarrow \Box^{*\,m+2}\bot. \tag{17}$$

So, by the Strong Löb Principle, $\mathsf{HA}^* \vdash \Box^{*\,n+2}\bot$. Quod non. Thus, the k_n take infinitely many values. We may conclude that A. ❑

References

[AB04] S.N. Artemov and L.D. Beklemishev. Provability logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd ed.*, volume 13, pages 229–403. Kluwer, Dordrecht, 2004.

[Art80] S.N. Artëmov. Arithmetically complete modal theories. In *Semiotika i Informatika,* 14, pages 115–133. VINITI, Moscow, 1980. In Russian. English translation in: *Amer. Math. Soc. Transl.* (2), 135: 39–54, 1987.

[Bee75] M. Beeson. The nonderivability in intuitionistic formal systems of theorems on the continuity of effective operations. *The Journal of Symbolic Logic*, 40:321–346, 1975.

[BM87] G. Boolos and V. McGee. The degree of the set of sentences of predicate provability logic that are true under every interpretation. *The Journal of Symbolic Logic*, 52(1):165–171, 1987.

[Boo93] G. Boolos. *The logic of provability.* Cambridge University Press, Cambridge, 1993.

[Bur00] W. Burr. Fragments of Heyting Arithmetic. *The Journal of Symbolic Logic*, 65(3):1223–1240, 2000.

[dJV96] D.H.J. de Jongh and A. Visser. Embeddings of Heyting algebras. In *[HHST96]*, pages 187–213, 1996.

[HHST96] W. Hodges, M. Hyland, C. Steinhorn, and J. Truss, editors. *Logic: from foundations to applications.* Clarendon Press, Oxford, 1996.

[Lif79] V. Lifschitz. CT_0 is stronger than CT_0! *Proceedings of the American Mathematical Society*, 73:101–106, 1979.

[McC91] D.C. McCarty. Incompleteness in intuitionistic metamathematics. *The Notre Dame Journal of Formal Logic*, 32:323–358, 1991.

[Mon87] F. Montagna. The predicate modal logic of provability. *Notre Dame Journal of Formal Logic*, 25:179–189, 1987.

[Oos90] J. van Oosten. Lifschitz' realizability. *The Journal of Symbolic Logic*, 55:805–821, 1990.

[Pli73] V. E. Plisko. On realizable predicative formulas. *Soviet Math. Dokl.*, 14:1420–1423, 1973.

[Pli77] V. E. Plisko. The nonarithmeticity of the class of realizable formulas. *Math. of USSR Izv.*, 11:453–471, 1977.

[Pli78] V. E. Plisko. Some variants of the notion of realizability for predicate formulas. *Math. of USSR Izv.*, 12:588–604, 1978.

[Pli83] V. E. Plisko. Absolute realizability of predicate formulas. *Math. of USSR Izv.*, 22:291–308, 1983.

[Pli90] V. E. Plisko. Modified realizability and predicate logic. *Abstracts of the All Union Conference in mathematical logic, Alma Ata, 1990*, 1990. Russian.

[Pli91] V. E. Plisko. Constructive formalization of the Tennenbaum theorem. *Mathematical Notes*, pages 950–957, 1991. Translated from *Mat. Zametki*, vol. 48, pp. 108–118, 1990.

[Pli93] V. E. Plisko. On arithmetic complexity of certain constructive logics. *Mathematical Notes*, pages 701–709, 1993. Translated from *Mat. Zametki*, vol. 52, pp. 94–104, 1992.

[Pli02] V. E. Plisko. Arithmetic complexity of the predicate logics of certain complete arithmetic theories. *Annals of Pure and Applied Logic*, 113:243–259, 2002.

[TvD88] A.S. Troelstra and D. van Dalen. *Constructivism in Mathematics, vol 1.* Studies in Logic and the Foundations of Mathematics, vol. 121. North Holland, Amsterdam, 1988.

[Var86] V.A. Vardanyan. Arithmetic comlexity of predicate logics of provability and their fragments. *Doklady Akad. Nauk SSSR*, 288(1):11–14, 1986. In Russian. English translation in *Soviet Mathematics Doklady* 33:569–572, 1986.

[Vis82] A. Visser. On the completeness principle. *Annals of Mathematical Logic*, 22:263–295, 1982.

[Vis99] A. Visser. Rules and Arithmetics. *The Notre Dame Journal of Formal Logic*, 40(1):116–140, 1999.

[Vis02] A. Visser. Substitutions of Σ^0_1-sentences: explorations between intuitionistic propositional logic and intuitionistic arithmetic. *Annals of Pure and Applied Logic*, 114:227–271, 2002.

[Vis03] A. Visser. No escape from Vardanyan's Theorem. Logic Group Preprint Series 220, Department of Philosophy, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, `http://preprints.phil.uu.nl/lgps/`, April 2003.

[Vis04] A. Visser. Categories of Theories and Interpretations. Logic Group Preprint Series 228, Department of Philosophy, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, `http://preprints.phil.uu.nl/lgps/`, April 2004.

[Vis05] A. Visser. Faith & Falsity: a study of faithful interpretations and false Σ^0_1-sentences. *Annals of Pure and Applied Logic*, 131:103–131, 2005.

[Yav97] R.E. Yavorsky. Logical schemes for first order theories. In *Springer LNCS (Yaroslavl'97 volume)*, volume 1234, pages 410–418, 1997.

A The Friedman Translation

In this appendix we give a generalization of Theorem 7.2. I do not know any application of the generalization, but it seems fairly natural.

The Friedman interpretation $(\cdot)^B$ is defined as follows. It sends atomic formulas α to $\alpha \vee B$. It commutes with all connectives including the quantifiers, with the exception of $\bot$ which is treated like an atomic formula. Note that the Friedman translation is not a relative translation in the sense of this paper.

Theorem A.1 *Suppose all the free variables in A are among $\vec{x}$. We have:*
$\mathsf{HA}^* + \mathsf{F}^K \vdash \forall \vec{x} \forall \vec{y} {\in} \delta_K \ (\vec{x} F_K \vec{y} \to (A^{\Box^*\bot}\vec{x} \leftrightarrow (A^{\Box^*\bot}\vec{y})^K))$.

Proof

The proof is by induction on A. We reason in $\mathsf{HA}^* + \mathsf{F}^K$.

Let A be an atomic formula α. Suppose $\vec{x} F_K \vec{y}$. Since F_K is an initial embedding, we have $\alpha\vec{x}$ iff $(\alpha\vec{y})^K$. By Theorem 7.2, we have $\Box^*\bot$ iff $(\Box^*\bot)^K$. So, we have:

$$\begin{array}{rcl} \alpha^{\Box^*\bot}\vec{x} & \leftrightarrow & \alpha\vec{x} \vee \Box^*\bot \\ & \leftrightarrow & (\alpha\vec{y})^K \vee (\Box^*\bot)^K \\ & \leftrightarrow & (\alpha\vec{y} \vee \Box^*\bot)^K \\ & \leftrightarrow & (\alpha^{\Box^*\bot}\vec{y})^K \end{array}$$

The cases of the propositional connectives are trivial.

Let $A\vec{x}$ be $\exists u\, Bu\vec{x}$. Suppose $\vec{x}F_K\vec{y}$. We treat the right to left direction. Suppose $(\exists v\, Bv\vec{y})^{\Box^*\bot,K}$. Say v witnesses this. We either have uF_Kv, for some u, or $\Box^*\bot$. In case $\Box^*\bot$, we are easily done. So assume uF_Kv. We have $u, \vec{x}F_Kv, \vec{y}$. By the induction hypothesis, we find $B^{\Box^*\bot}u\vec{x}$ and, hence $(\exists u\, Bu\vec{x})^{\Box^*\bot}$. The other direction is easier.

The case of the universal quantifier is similar. ❑