Institute for Language, Logic and Information

FEASIBLE INTERPRETABILITY

Rineke Verbrugge

ITLI Prepublication Series
for Mathematical Logic and Foundations ML-91-11

University of Amsterdam

The ITLI Prepublication Series

1986 86-01 The Institute of Language, Logic and Information
86-02 Peter van Emde Boas A Semantical Model for Integration and Modularization of Rules
86-03 Johan van Benthem Categorial Grammar and Lambda Calculus
86-04 Reinhard Muskens A Relational Formulation of the Theory of Types
86-05 Kenneth A. Bowen, Dick de Jongh Some Complete Logics for Branched Time, Part I Well-founded Time, Forward looking Operators
86-06 Johan van Benthem Logical Syntax
1987 87-01 Jeroen Groenendijk, Martin Stokhof Type shifting Rules and the Semantics of Interrogatives
87-02 Renate Bartsch Frame Representations and Discourse Representations
87-03 Jan Willem Klop, Roel de Vrijer Unique Normal Forms for Lambda Calculus with Surjective Pairing
87-04 Johan van Benthem Polyadic quantifiers
87-05 Víctor Sánchez Valencia Traditional Logicians and de Morgan's Example
87-06 Eleonore Oversteegen Temporal Adverbials in the Two Track Theory of Time
87-07 Johan van Benthem Categorial Grammar and Type Theory
87-08 Renate Bartsch The Construction of Properties under Perspectives
87-09 Herman Hendriks Type Change in Semantics: The Scope of Quantification and Coordination
1988 LP-88-01 Michiel van Lambalgen *Logic, Semantics and Philosophy of Language:* Algorithmic Information Theory
LP-88-02 Yde Venema Expressiveness and Completeness of an Interval Tense Logic
LP-88-03 Year Report 1987
LP-88-04 Reinhard Muskens Going partial in Montague Grammar
LP-88-05 Johan van Benthem Logical Constants across Varying Types
LP-88-06 Johan van Benthem Semantic Parallels in Natural Language and Computation
LP-88-07 Renate Bartsch Tenses, Aspects, and their Scopes in Discourse
LP-88-08 Jeroen Groenendijk, Martin Stokhof Context and Information in Dynamic Semantics
LP-88-09 Theo M.V. Janssen A mathematical model for the CAT framework of Eurotra
LP-88-10 Anneke Kleppe A Blissymbolics Translation Program
ML-88-01 Jaap van Oosten *Mathematical Logic and Foundations:* Lifschitz' Realizabiility
ML-88-02 M.D.G. Swaen The Arithmetical Fragment of Martin Löf's Type Theories with weak Σ-elimination
ML-88-03 Dick de Jongh, Frank Veltman Provability Logics for Relative Interpretability
ML-88-04 A.S. Troelstra On the Early History of Intuitionistic Logic
ML-88-05 A.S. Troelstra Remarks on Intuitionism and the Philosophy of Mathematics
CT-88-01 Ming Li, Paul M.B.Vitanyi *Computation and Complexity Theory:* Two Decades of Applied Kolmogorov Complexity
CT-88-02 Michiel H.M. Smid General Lower Bounds for the Partitioning of Range Trees
CT-88-03 Michiel H.M. Smid, Mark H. Overmars, Leen Torenvliet, Peter van Emde Boas
Maintaining Multiple Representations of Dynamic Data Structures
CT-88-04 Dick de Jongh, Lex Hendriks, Gerard R. Renardel de Lavalette Computations in Fragments of Intuitionistic Propositional Logic
CT-88-05 Peter van Emde Boas Machine Models and Simulations (revised version)
CT-88-06 Michiel H.M. Smid A Data Structure for the Union-find Problem having good Single-Operation Complexity
CT-88-07 Johan van Benthem Time, Logic and Computation
CT-88-08 Michiel H.M. Smid, Mark H. Overmars, Leen Torenvliet, Peter van Emde Boas Multiple Representations of Dynamic Data Structures
CT-88-09 Theo M.V. Janssen Towards a Universal Parsing Algorithm for Functional Grammar
CT-88-10 Edith Spaan, Leen Torenvliet, Peter van Emde Boas Nondeterminism, Fairness and a Fundamental Analogy
CT-88-11 Sieger van Denneheuvel, Peter van Emde Boas Towards implementing RL
X-88-01 Marc Jumelet *Other prepublications:* On Solovay's Completeness Theorem
1989 LP-89-01 Johan van Benthem *Logic, Semantics and Philosophy of Language:* The Fine-Structure of Categorial Semantics
LP-89-02 Jeroen Groenendijk, Martin Stokhof Dynamic Predicate Logic, towards a compositional, non-representational semantics of discourse
LP-89-03 Yde Venema Two-dimensional Modal Logics for Relation Algebras and Temporal Logic of Intervals
LP-89-04 Johan van Benthem Language in Action
LP-89-05 Johan van Benthem Modal Logic as a Theory of Information
LP-89-06 Andreja Prijatelj Intensional Lambek Calculi: Theory and Application
LP-89-07 Heinrich Wansing The Adequacy Problem for Sequential Propositional Logic
LP-89-08 Víctor Sánchez Valencia Peirce's Propositional Logic: From Algebra to Graphs
LP-89-09 Zhisheng Huang Dependency of Belief in Distributed Systems
ML-89-01 Dick de Jongh, Albert Visser *Mathematical Logic and Foundations:* Explicit Fixed Points for Interpretability Logic
ML-89-02 Roel de Vrijer Extending the Lambda Calculus with Surjective Pairing is conservative
ML-89-03 Dick de Jongh, Franco Montagna Rosser Orderings and Free Variables
ML-89-04 Dick de Jongh, Marc Jumelet, Franco Montagna On the Proof of Solovay's Theorem
ML-89-05 Rineke Verbrugge Σ-completeness and Bounded Arithmetic
ML-89-06 Michiel van Lambalgen The Axiomatization of Randomness
ML-89-07 Dirk Roorda Elementary Inductive Definitions in HA: from Strictly Positive towards Monotone
ML-89-08 Dirk Roorda Investigations into Classical Linear Logic
ML-89-09 Alessandra Carbone Provable Fixed points in $I\Delta_0+\Omega_1$
CT-89-01 Michiel H.M. Smid *Computation and Complexity Theory:* Dynamic Deferred Data Structures
CT-89-02 Peter van Emde Boas Machine Models and Simulations
CT-89-03 Ming Li, Herman Neuféglise, Leen Torenvliet, Peter van Emde Boas On Space Efficient Simulations
CT-89-04 Harry Buhrman, Leen Torenvliet A Comparison of Reductions on Nondeterministic Space
CT-89-05 Pieter H. Hartel, Michiel H.M. Smid, Leen Torenvliet, Willem G. Vree A Parallel Functional Implementation of Range Queries
CT-89-06 H.W. Lenstra, Jr. Finding Isomorphisms between Finite Fields
CT-89-07 Ming Li, Paul M.B. Vitanyi A Theory of Learning Simple Concepts under Simple Distributions and Average Case Complexity for the Universal Distribution (Prel. Version)
CT-89-08 Harry Buhrman, Steven Homer, Leen Torenvliet Honest Reductions, Completeness and Nondeterminstic Complexity Classes
CT-89-09 Harry Buhrman, Edith Spaan, Leen Torenvliet On Adaptive Resource Bounded Computations
CT-89-10 Sieger van Denneheuvel The Rule Language RL/1
CT-89-11 Zhisheng Huang, Sieger van Denneheuvel, Peter van Emde Boas Towards Functional Classification of Recursive Query Processing
X-89-01 Marianne Kalsbeek *Other Prepublications:* An Orey Sentence for Predicative Arithmetic
X-89-02 G. Wagemakers New Foundations: a Survey of Quine's Set Theory
X-89-03 A.S. Troelstra Index of the Heyting Nachlass
X-89-04 Jeroen Groenendijk, Martin Stokhof Dynamic Montague Grammar, a first sketch
X-89-05 Maarten de Rijke The Modal Theory of Inequality
X-89-06 Peter van Emde Boas Een Relationele Semantiek voor Conceptueel Modelleren: Het RL-project
1990 *Logic, Semantics and Philosophy of Language*
LP-90-01 Jaap van der Does A Generalized Quantifier Logic for Naked Infinitives
LP-90-02 Jeroen Groenendijk, Martin Stokhof Dynamic Montague Grammar
LP-90-03 Renate Bartsch Concept Formation and Concept Composition
LP-90-04 Aarne Ranta Intuitionistic Categorial Grammar
LP-90-05 Patrick Blackburn Nominal Tense Logic
LP-90-06 Gennaro Chierchia The Variablity of Impersonal Subjects
LP-90-07 Gennaro Chierchia Anaphora and Dynamic Logic
LP-90-08 Herman Hendriks Flexible Montague Grammar
LP-90-09 Paul Dekker The Scope of Negation in Discourse, towards a flexible dynamic Montague grammar
LP-90-10 Theo M.V. Janssen Models for Discourse Markers
LP-90-11 Johan van Benthem General Dynamics
LP-90-12 Serge Lapierre A Functional Partial Semantics for Intensional Logic

Instituut voor Taal, Logica en Informatie
Institute for Language, Logic and Information

Faculteit der Wiskunde en Informatica
(Department of Mathematics and Computer Science)
Plantage Muidergracht 24
1018TV Amsterdam

Faculteit der Wijsbegeerte
(Department of Philosophy)
Nieuwe Doelenstraat 15
1012CP Amsterdam

FEASIBLE INTERPRETABILITY

Rineke Verbrugge
Department of Mathematics and Computer Science
University of Amsterdam

ITLI Prepublications
for Mathematical Logic and Foundations
ISSN 0924-2090

Received October 1991

Research supported by the
Netherlands Organization for Scientific Research (NWO)

Feasible interpretability

Rineke Verbrugge
Faculty of Mathematics and Computer Science
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
the Netherlands

October 15, 1991

Abstract

In PA, or even in $I\Delta_0 + EXP$, we can define the concept of feasible interpretability. Informally stated, U feasibly interprets V iff:

> for some interpretation, U proves the interpretations of all axioms of V by proofs with Gödel numbers of length polynomial in the length of the Gödel numbers of those axioms.

Here both U and V are Σ_1^b-axiomatized theories.

Many interpretations encountered in everyday mathematics (e.g. the interpretation of Peano arithmetic into ZF) are feasible. However, by fixed point constructions we can find theories that are interpretable in PA in the usual sense but not by a feasible interpretation. By making polynomial analogs of the usual proofs, we show that the bimodal interpretability logic ILM is sound for feasible interpretability over the base theory PA. Here, $A \rhd B$ is translated as $PA + A^* \rhd_f PA + B^*$, where * is the translation. Moreover, we can prove in PA a polynomial version of Orey's theorem for feasible interpretability. This paves the way for a polynomial adaptation of Berarducci's proof of arithmetical completeness of ILM with respect to PA. Thus, we show that ILM is arithmetically sound and complete with respect to feasible interpretability over PA.

1 Introduction

In this paper, we investigate a new concept of interpretability – we call it feasible interpretability – in which the complexity of the interpretation is bounded in a certain way. The concept was invented by Albert Visser, who called it effective interpretability in his paper [Vi].

In order to define this concept, we first review the usual definition of interpretability.

Let U, V be two Σ_1^b-axiomatized theories, where V is axiomatized by the Σ_1^b-formula α_V. An interpretation K of V into U is given as usual by a formula $\delta(x)$ of L_U defining the universe, and a function from the relation and function symbols of L_V to formulas of L_U, respecting the original arities. In the sequel we take the image of $=$ to be $=$, though this is not essential for the results. We can extend K in the usual way to map all formulas

φ of L_V into formulas φ^K of L_U; in fact we can, in an intensionally correct way, Δ_1^b-define in $I\Delta_0 + \Omega_1$ a function K corresponding to this mapping. For ease of reading, we will write a^K even if a is a Gödel number. Thus $U \rhd V$ can be defined in $I\Delta_0 + \Omega_1$ as follows:

$$I\Delta_0 + \Omega_1 \vdash U \rhd V \leftrightarrow \exists K(\text{“}K \mathit{is\ an\ interpretation}\text{”} \wedge \forall a(\alpha_V(a) \rightarrow \exists p Prf_U(p, a^K))).$$

Similarly, we would like to define a concept of *feasible interpretability*, given half-formally as

$$U \rhd_f V \leftrightarrow \exists K \exists P(\text{“}K \mathit{is\ an\ interpretation\ and}\ P \mathit{is\ a\ polynomial}\text{”} \wedge \\ \forall a(\alpha_V(a) \rightarrow \exists p(\text{“}|p| \leq P(|a|)\text{”} \wedge Prf_U(p, a^K)))) \quad (1)$$

If we want to formalize this concept, we need an evaluation function for coded polynomials and we need to be able to prove that the *exp* of this function is total. We remind the reader that *exp*(the values of polynomials in $|x|$) corresponds to the values of #-terms in x, where $x\#y = exp(|x| \cdot |y|)$ as defined in Buss[86]. Thus, since there is an evaluation function for formalized terms containing # that is provably total in $I\Delta_0 + EXP$, we see that the formalization of feasible interpretability can be carried out in $I\Delta_0 + EXP$. We will not carry out the details, and for ease of reading we will keep using the half-formal definition(1).

However, it is clear that the formula $U \rhd_f V$ is Σ_2^0. As we know that, for reasonable theories U extending PA, $\{A | U \rhd U + A\}$ is a Π_2^0-complete predicate, it would be interesting to find out whether $\{A \mid U \rhd_f U + A\}$ is Σ_2^0-complete. We haven't yet found the answer to this question.

In [Vi], Visser gave proof sketches to show that ILM is arithmetically sound with respect to feasible interpretability over PA. Moreover, he gave an Orey-Hájek like characterization for feasible interpretability over PA^*, where PA^* is defined as follows:

> C is an axiom of PA^* iff C is the conjunction of the first n axioms of PA for some n.

He then surmised that, using this characterization, Berarducci's arguments from [Be 90] could be adapted to show that ILM is the modal interpretability logic for feasible interpretability over PA^*.

In this paper, we show that ILM is indeed arithmetically sound and complete with respect to feasible interpretability over PA itself.

The rest of the paper is organized as follows. In section 2, we show that some well-known interpretations from the contexts of set theory and bounded arithmetic are feasible. For the subsequent sections, the horizon is narrowed down to Peano Arithmetic. Thus we prove in section 3 and section 5 that ILM is exactly the modal interpretability logic for feasible interpretability over PA. Section 4, meanwhile, gives two counterexamples to show that, for reasonable theories U extending PA, feasible interpretability over U is a definitely stricter concept than normal interpretability.

2 Feasible interpretations in various settings

For an intuitive introduction to feasible interpretability, it is useful to define feasible interpretability also for settings other than arithmetic. The informal definition is as follows.

$U \rhd_f V$ if and only if there is an interpretation K of V into U which is feasible, i.e. for which there is a polynomial P such that for all axioms φ of V, there is a proof of length $\leq P(|\varphi|)$ in U of φ^K

Here $|\varphi|$ denotes the length of φ. In this section, we look at some well-known interpretations from different settings and show that they are feasible. As a first remark, it is clear that every interpretation of a finitely axiomatized theory into some other theory is feasible: a constant polynomial, namely the maximum of the lengths of the proofs of the interpreted axioms, suffices. We first prove an easy lemma which can be used to show that many well-known interpretations are feasible.

Remark 2.1 Of course the definitions of $|\varphi|$ and of the lengths of proofs depend on the setting. For example, it is not always convenient to define $|\varphi|$ as "the length of the binary expression for the Gödel number of φ".

However, we have to keep in mind that a few conditions on the definition of the lengths of formulas and proofs are necessary to make lemma 2.2 applicable.

The length of formulas should be defined in such a way that the following conditions hold:

1. $|\neg\psi| \geq |\psi| + 1$,
2. $|\psi \circ \chi| \geq |\psi| + |\chi| + 1$ for $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$,
3. $|Qx\psi| \geq |\psi| + 1$ for $Q \in \{\forall, \exists\}$, and
4. for all formulas φ, $|\varphi| \geq 2$.

The last of these conditions is not necessary, but it just simplifies the computations by allowing us to work with polynomials $P(n)$ of the form n^d only.

Moreover, we suppose that the proof system and the corresponding length of a proof is defined in such a way that applications of $\wedge$-rules and Modus Ponens do not make the proofs explode to an inordinate length; e.g. we suppose that we do not use a tableau system or a sequent calculus without the cut rule. A sufficient condition is the following.

There is a constant c such that the following conditions hold:

1. if we have a proof of of length l_A of the formula A, and a proof of length $l_{A\rightarrow B}$ of $A \rightarrow B$, then there is a proof of length $\leq l_A + l_{A\rightarrow B} + |B|^c$ of the formula B; and
2. if we have a proof of length l_A of A, a proof of length l_B of B and a proof of length $l_{A\wedge B\rightarrow C}$ of $A \wedge B \rightarrow C$, then we have a proof of length $\leq l_A + l_B + l_{A\wedge B\rightarrow C} + |C|^c$ of the formula C.

Lemma 2.2 *Let L be a language and U a theory satisfying the conditions in Remark 2.1. Let F be a function from L into L_U such that*

there is a polynomial P such that for all $\varphi \in L$, $|F(\varphi)| \leq P(|\varphi|)$.

Moreover, suppose that U proves the following by proofs of length $\leq P(|\varphi|)$, resp. $\leq P(|\neg\psi|)$, resp. $\leq P(|\psi \circ \chi|)$, resp. $\leq P(|Qx\psi|)$:

1. *$F(\varphi)$ for all atomic $\varphi \in L$;*

2. $F(\psi) \rightarrow F(\neg\psi)$ for all $\psi \in L$;

3. $F(\psi) \wedge F(\chi) \rightarrow F(\psi \circ \chi)$ for all $\psi, \chi \in L$ and $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$;

4. $F(\psi) \rightarrow F(Qx\psi)$ for all $\psi \in L$ and $Q \in \{\forall, \exists\}$.

Then there is a polynomial R such that for all $\varphi \in L$, $U \vdash \varphi$ by a proof of length $\leq R(|\varphi|)$.

Proof. Take a constant $d \geq 2$ such that

1. for all $n \geq 2$, $P(n) \leq n^d$ and
2. for all $\varphi \in L$, $|F(\varphi)|^c \leq |\varphi|^d$, where c is as in Remark 2.1 in the condition on the length of proofs.

Define the polynomial $R(n) := n^{2d}$. We will prove by induction on the construction of φ that for all $\varphi \in L$, $U \vdash F(\varphi)$ by a proof of length $\leq R(|\varphi|)$.

Basic step By the assumption we know that for atomic formulas φ, $U \vdash F(\varphi)$ by a proof of length $\leq P(|\varphi|)$. But by definition of d, $P(|\varphi|) \leq |\varphi|^d \leq |\varphi|^{2d}$.

$\neg$-step Suppose as induction hypothesis that $U \vdash F(\psi)$ by a proof of length $\leq |\psi|^{2d}$. By assumption, $U \vdash F(\psi) \rightarrow F(\neg\psi)$ by a proof of length $\leq P(|\neg\psi|) \leq |\neg\psi|^d$ (where the last inequality holds because of clause 1 of the definition of d). Therefore by the first clause in the condition on the length of proofs in Remark 2.1, we have $U \vdash F(\neg\psi)$ by a proof of length $\leq |\psi|^{2d} + |\neg\psi|^d + |F(\neg\psi)|^c \leq |\psi|^{2d} + |\neg\psi|^d + |\neg\psi|^d$ (where the last inequality holds by clause 2 of the definition of d). Since we assume that $|\neg\psi| \geq |\psi| + 1$, we have $|\psi|^{2d} + |\neg\psi|^d + |\neg\psi|^d \leq |\neg\psi|^{2d}$ by an easy computation using the binomial theorem and the fact that $d \geq 2$. The quantifier steps are analogous to the $\neg$-step, so we leave them to the reader.

Connective step Let $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$. Suppose as induction hypothesis that $U \vdash F(\psi)$ by a proof of length $\leq |\psi|^{2d}$, and $U \vdash F(\chi)$ by a proof of length $\leq |\chi|^{2d}$. By assumption, $U \vdash F(\psi) \wedge F(\chi) \rightarrow F(\psi \circ \chi)$ by a proof of length $\leq P(|\psi \circ \chi|) \leq |\psi \circ \chi|^d$.

The second clause in the condition on the length of proofs in Remark 2.1 now implies that $U \vdash F(\psi \circ \chi)$ by a proof of length $\leq |\psi|^{2d} + |\chi|^{2d} + |\psi \circ \chi|^d + |F(\psi \circ \chi)|^c \leq |\psi|^{2d} + |\chi|^{2d} + |\psi \circ \chi|^d + |\psi \circ \chi|^d$ (where the last inequality holds by clause 2 in the definition of d).

Since we assume that $|\psi \circ \chi| \geq |\psi| + |\chi| + 1$, we can again use the binomial theorem to show that $|\psi|^{2d} + |\chi|^{2d} + |\psi \circ \chi|^d + |\psi \circ \chi|^d \leq |\psi \circ \chi|^{2d}$, as desired.

QED

Remark 2.3 When we want to prove that some interpretation K of V into U is feasible, we can often use Lemma 2.2. Suppose all axioms of V have the form $\Phi(\psi)$, where Φ is a formula scheme. The feature we need in order to apply Lemma 2.2 is the fact that both $|\Phi(\psi)|$ and $|\psi^K|$ are polynomial in $|\psi|$.

As a first example, in which we do not yet need lemma 2.2, we will show that the usual interpretation of $\mathrm{I}\Delta_0 + \Omega_1$ into $\mathrm{I}\Delta_0$ by a cut is feasible.

Theorem 2.4 *$I\Delta_0 \rhd_f I\Delta_0 + \Omega_1$ by a cut.*

Proof. Proof. Let J be a cut constructed by Solovay's methods such that $\mathrm{I}\Delta_0$ proves that J is a cut closed under $+, \cdot$, and ω_1. Define φ^J to be the formula φ with all quantifiers restricted to J. It is well-known that J is an interpretation of $\mathrm{I}\Delta_0 + \Omega_1$ into $\mathrm{I}\Delta_0$; so to show that it is a feasible interpretation, it suffices to find a polynomial P such that for all Δ_1^0-formulas φ, the following holds by proofs of length $\leq P(|\ulcorner\varphi\urcorner|)$:

$$\mathrm{I}\Delta_0 \vdash [\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(Sx)) \rightarrow \forall x \varphi(x)]^J$$

First, it is easy to see that there is a polynomial P_1 such that for all Δ_1^0-formulas φ, $\mathrm{I}\Delta_0 \vdash J(a) \rightarrow (\varphi(a) \leftrightarrow \varphi(a)^J)$ and $\mathrm{I}\Delta_0 \vdash \forall x\varphi \rightarrow (\forall x \varphi)^J$ by proofs of length $\leq P_1(|\ulcorner\varphi\urcorner|)$. Second, there is a polynomial P_2 such that for all Δ_1^0-formulas φ, the following holds by proofs of length $\leq P_2(|\ulcorner\varphi\urcorner|)$:

$$\mathrm{I}\Delta_0 \vdash \forall a\, [\varphi(0) \wedge \forall x \leq a\, (\varphi(x) \rightarrow \varphi(Sx)) \rightarrow \forall x \leq a\, \varphi(x)]$$

In fact one uses only the induction axiom for $\forall x \leq a\, \varphi(x)$, the fact that $\forall a \forall x(Sx \leq a \rightarrow x \leq a)$, and some predicate logic. Combining P_1 with P_2, we then find a polynomial P_3 such that for all Δ_1^0-formulas φ, the following holds by proofs of length $\leq P_3(|\ulcorner\varphi\urcorner|)$:

$$\mathrm{I}\Delta_0 \vdash (\forall a[\varphi(0) \wedge \forall x \leq a(\varphi(x) \rightarrow \varphi(Sx)) \rightarrow \forall x \leq a\, \varphi(x)])^J$$

Now it is easy to find a polynomial P from P_3 such that for all Δ_1^0-formulas φ, the following holds by proofs of length $\leq P(|\ulcorner\varphi\urcorner|)$:

$$\mathrm{I}\Delta_0 \vdash [\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(Sx)) \rightarrow \forall x \varphi(x)]^J$$

We use only the fact that $\forall a(a \leq a)$ and some predicate logic. Thus, J is a feasible interpretation of $\mathrm{I}\Delta_0 + \Omega_1$ into $\mathrm{I}\Delta_0$. QED

Next, we will prove that the usual interpretation of $ZF + \mathbf{V} = \mathbf{L}$ into ZF is feasible. Because ZF consists of a finite list of axioms plus the schemata of separation and replacement, we can restrict our attention to feasibly proving these schemata relativized to the universe $\mathbf{L}$ of constructible sets. We will first prove that the schema of separation relativized to $\mathbf{L}$ follows feasibly from the reflection theorem for $\mathbf{L}$, and then give a feasible proof of the reflection theorem itself. We will try to follow the elegant proof in terms of closed unbounded collections, which unfortunately becomes much less elegant when forced into the straightjacket of the calculation of lengths. We will not stray far from the straightforward presentation given in [Ku 80], where all details about the constructible universe that we omit here can be found. The length $|\varphi|$ of a formula φ of ZF is defined as the number of appearances of symbols in φ; without loss of generality, we can take the length of all variables to be 1. Likewise, we define the length of a proof in ZF to be the total number of symbols appearing in the proof. In the following lemmas, quantifiers in greek letters range over the ordinals, while those in roman letters range over all sets.

The next lemma corresponds to lemma IV.2.5 of [Ku 80].

Lemma 2.5 *ZF proves the following by proofs of length polynomial in* $|\varphi|$*:*

$$\begin{aligned} \forall z, \overline{v} \in \mathbf{L} \quad & \{x \in z \mid \varphi^{\mathbf{L}}(x, z, \overline{v})\} \in \mathbf{L} \rightarrow \\ & \forall z, \overline{v} \in \mathbf{L}\, \exists y \in \mathbf{L}\, [x \in y \leftrightarrow x \in z \wedge \varphi^{\mathbf{L}}(x, z, \overline{v})] \end{aligned}$$

Proof. Straightforward; we do not even need the fact that $\mathbf{L}$ is transitive. Note that by absoluteness of atomic formulas for $\mathbf{L}, \mathbf{V}$, the succedent is feasibly equivalent to the comprehension schema for φ, relativized to $\mathbf{L}$. QED

The following lemma corresponds to a part of lemma VI.2.1 of [Ku 80].

Lemma 2.6 *ZF proves the following by proofs of length polynomial in* $|\varphi|$*:*

$$\begin{aligned} \forall \alpha\, \exists \beta\, \forall z, x, \overline{v} \in \mathbf{L}_\beta \quad & [\varphi^{\mathbf{L}}(x, z, \overline{v}) \leftrightarrow \varphi^{\mathbf{L}_\beta}(x, z, \overline{v})] \rightarrow \\ & \forall z, \overline{v} \in \mathbf{L}\, \{x \in z \mid \varphi^{\mathbf{L}}(x, z, \overline{v})\} \in \mathbf{L} \end{aligned}$$

Proof. It is easy to see that the usual proof in *ZF* is feasible: suppose

1. $\forall \alpha\, \exists \beta\, \forall z, x, \overline{v} \in \mathbf{L}_\beta\, [\varphi^{\mathbf{L}}(x, z, \overline{v}) \leftrightarrow \varphi^{\mathbf{L}_\beta}(x, z, \overline{v})]$ and
2. $z, \overline{v} \in \mathbf{L}$

From 2 it follows that there is an α such that $z, \overline{v} \in \mathbf{L}_\alpha$. Now let $\beta > \alpha$ be such that $\forall x \in \mathbf{L}_\beta\, [\varphi^{\mathbf{L}}(x, z, \overline{v}) \leftrightarrow \varphi^{\mathbf{L}_\beta}(x, z, \overline{v})]$. Then, using the fact that $\mathbf{L}$ is transitive and that $x \in z$ is absolute for $\mathbf{L}_\beta, \mathbf{L}$, we find that

$$\{x \in z \mid \varphi^{\mathbf{L}}(x, z, \overline{v})\} = \{x \in \mathbf{L}_\beta \mid (x \in z \wedge \varphi(x, z, \overline{v}))^{\mathbf{L}_\beta}\} \in Def(\mathbf{L}_\beta) = \mathbf{L}_{\beta+1},$$

so $\{x \in z \mid \varphi^{\mathbf{L}}(x, z, \overline{v})\} \in \mathbf{L}$. QED

From lemma 2.5 and lemma 2.6, we conclude that in order to feasibly prove the comprehension schema, we only need polynomial length proofs of

$$\forall \alpha\, \exists \beta\, \forall z, x, \overline{v} \in \mathbf{L}_\beta\, [\varphi^{\mathbf{L}}(x, z, \overline{v}) \leftrightarrow \varphi^{\mathbf{L}_\beta}(x, z, \overline{v})].$$

For a proof of this reflection theorem, we need a few more definitions.

Definition 2.7 A collection $\mathcal{C}$ of ordinals is

- *unbounded* iff $\forall \alpha\, \exists \beta > \alpha\, (\beta \in \mathcal{C})$;
- *closed* iff $\forall \alpha\, (\alpha \neq \emptyset \wedge a \subseteq \mathcal{C} \rightarrow \sup\, a \in \mathcal{C})$;
- *closed unbounded* (c.u.b.)] iff $\mathcal{C}$ is both closed and unbounded.

Lemma 2.8 $ZF \vdash$ *"If* $\mathcal{C}$ *and* $\mathcal{D}$ *are c.u.b., then* $\mathcal{C} \cap \mathcal{D}$ *is c.u.b. as well"*

Proof. An easy application of lemma II.6.8 of [Ku 80]. QED

Definition 2.9 A collection $\mathcal{C}$ of ordinals is *closed unbounded for* φ iff

1. $\mathcal{C}$ is closed unbounded, and

2. $\mathcal{C}$ consists of ordinals α such that $\mathbf{L}_\alpha$ reflects φ, i.e.
$\forall\alpha\,(\alpha \in \mathcal{C} \rightarrow \forall\overline{v}\in\mathbf{L}_\alpha\,[\varphi^{\mathbf{L}}(\overline{v}) \leftrightarrow \varphi^{\mathbf{L}_\alpha}(\overline{v})])$

Suppose φ is a formula and $\mathcal{D}$ is a first-order definable collection of ordinals. Using definition 2.7, we are able to construct new first-order formulas $CUB_{\mathcal{D}}$, $CUB_{\mathcal{D},\varphi}$ and REF_φ with the following intended meanings:

1. $CUB_{\mathcal{D}}$:= "$\mathcal{D}$ is closed unbounded"

2. $CUB_{\mathcal{D},\varphi}$:="$\mathcal{D}$ is closed unbounded for φ"

3. REF_φ:= " there is some collection of ordinals that is closed unbounded for φ"

The next lemma roughly corresponds to theorem IV.7.5 of [Ku 80].

Lemma 2.10 (Reflection theorem) *ZF proves the following by proofs of length polynomial in $|\varphi|$:*

$$\forall\alpha\,\exists\beta\,\forall z,x,\overline{v}\in\mathbf{L}_\beta\,[\varphi^{\mathbf{L}}(x,z,\overline{v}) \leftrightarrow \varphi^{\mathbf{L}_\beta}(x,z,\overline{v})]$$

Proof. First we note that ZF proves $(\alpha < \beta \rightarrow \mathbf{L}_\alpha \subseteq \mathbf{L}_\beta)$, "if γ is a limit ordinal, then $\mathbf{L}_\gamma = \bigcup_{\alpha<\gamma}\mathbf{L}_\alpha$" and $\mathbf{L} = \bigcup_{\alpha\in OR}\mathbf{L}_\alpha$.

We will prove the reflection theorem by induction on the construction of φ. A straightforward application of lemma 2.2 implies that for the reflection theorem to have a proof of length polynomial in $|\varphi|$, it is sufficient to find a polynomial bounding the lengths of the induction steps. Thus, we need to find a polynomial P such that by proofs of length $\leq P(|\varphi|)$, resp. $\leq P(|\neg\psi|)$, resp. $\leq P(|\psi\circ\chi)$, resp. $\leq P(|Qx\psi|)$, ZF proves the following:

1. for atomic φ:
$$\forall\alpha\,\exists\beta{>}\alpha\,\forall z,x\in\mathbf{L}_\beta\,[\varphi^{\mathbf{L}}(x,z) \leftrightarrow \varphi^{\mathbf{L}_\beta}(x,z)] \wedge CUB_{OR,\varphi}$$

2. the $\neg$-step:
$$\begin{aligned}&\forall\alpha\,\exists\beta{>}\alpha\,\forall\overline{v}\in\mathbf{L}_\beta\,[\psi^{\mathbf{L}}(\overline{v}) \leftrightarrow \psi^{\mathbf{L}_\beta}(\overline{v})] \wedge REF_\psi \quad \rightarrow\\ &\forall\alpha\,\exists\beta{>}\alpha\,\forall\overline{v}\in\mathbf{L}_\beta\,[\neg\psi^{\mathbf{L}}(\overline{v}) \leftrightarrow \neg\psi^{\mathbf{L}_\beta}(\overline{v})] \wedge REF_{\neg\psi}\end{aligned}$$

3. the connective step, where $\circ \in \{\wedge,\vee,\rightarrow,\leftrightarrow\}$:
$$\begin{aligned}&\forall\alpha\,\exists\beta{>}\alpha\,\forall\overline{v}\in\mathbf{L}_\beta\,[\psi^{\mathbf{L}}(\overline{v}) \leftrightarrow \psi^{\mathbf{L}_\beta}(\overline{v})] \wedge REF_\psi\wedge\\ &\forall\alpha\,\exists\beta{>}\alpha\,\forall\overline{w}\in\mathbf{L}_\beta\,[\chi^{\mathbf{L}}(\overline{w}) \leftrightarrow \chi^{\mathbf{L}_\beta}(\overline{w})] \wedge REF_\chi \quad \rightarrow\\ &\forall\alpha\,\exists\beta{>}\alpha\,\forall\overline{v}\in\mathbf{L}_\beta\,[\psi^{\mathbf{L}}\circ\chi^{\mathbf{L}}(\overline{v},\overline{w}) \leftrightarrow \psi^{\mathbf{L}_\beta}\circ\chi^{\mathbf{L}_\beta}(\overline{v},\overline{w})] \wedge REF_{\psi\circ\chi}\end{aligned}$$

4. the quantifier step, where $Q \in \{\exists,\forall\}$:
$$\begin{aligned}&\forall\alpha\,\exists\beta{>}\alpha\,\forall z,\overline{v}\in\mathbf{L}_\beta\,[\psi^{\mathbf{L}}(z,\overline{v}) \leftrightarrow \psi^{\mathbf{L}_\beta}(z,\overline{v})] \wedge REF_\psi \quad \rightarrow\\ &\forall\alpha\,\exists\beta{>}\alpha\,\forall\overline{v}\in\mathbf{L}_\beta\,[Qz\in\mathbf{L}\,\psi^{\mathbf{L}}(z,\overline{v}) \leftrightarrow Qz\in\mathbf{L}_\beta\,\psi^{\mathbf{L}_\beta}(z,\overline{v})] \quad \wedge REF_{Qz\psi}\end{aligned}$$

Finding polynomials bounding the lengths of the proofs of 1, 2 and 3 is very easy: we can use the feasibly provable fact that atomic formulas are absolute for any $\mathbf{L}_\alpha, \mathbf{L}$, some propositional reasoning independent on the specific ψ, χ, and an application of lemma 2.8 for step 3. We will show how the proofs of the $\exists$-case in step 4 can be bounded by a polynomial; a bound for the $\forall$-step then follows by some uses of the bounds for the $\neg$-step and the $\exists$-step.

Define

$$\mathcal{D} := \{\beta \mid \forall \overline{v} \in \mathbf{L}_\beta\, [\exists z \in \mathbf{L}\, \psi^{\mathbf{L}}(z, \overline{v}) \to \exists z \in \mathbf{L}_\beta\, \psi^{\mathbf{L}}(z, \overline{v})]\}.$$

It is easy to see that ZF proves the following by proofs of length polynomial in $|\exists z \psi|$:

$$\begin{aligned}&\forall \alpha\, \exists \beta > \alpha\, \forall z, \overline{v} \in \mathbf{L}_\beta\, [\psi^{\mathbf{L}}(z, \overline{v}) \leftrightarrow \psi^{\mathbf{L}_\beta}(z, \overline{v})] \wedge CUB_{\mathcal{C}_\psi, \psi} \wedge CUB_{\mathcal{D}} \to \\ &\forall \alpha\, \exists \beta > \alpha\, \forall \overline{v} \in \mathbf{L}_\beta\, [\exists z \in \mathbf{L}\, \psi^{\mathbf{L}}(z, \overline{v}) \leftrightarrow \exists z \in \mathbf{L}_\beta\, \psi^{\mathbf{L}}(z, \overline{v})] \wedge CUB_{\mathcal{C}_\psi \cap \mathcal{D}, \exists z \psi}\end{aligned}$$

In fact, we only use lemma 2.8 and the fact that $\forall \beta\, (\mathbf{L}_\beta \subseteq \mathbf{L})$. Thus we need to find a polynomial P such that $ZF \vdash CUB_{\mathcal{D}}$ by a proof of length $\leq P(|\exists z \psi|)$. Immediately from the definition, it is clear that $ZF \vdash$ "$\mathcal{D}$ is closed" by a proof of length polynomial in $|\exists z \psi|$. Thus, it suffices to show by a proof of length polynomial in $|\exists z \psi|$ that ZF proves that $\mathcal{D}$ is unbounded, that is:

$$\forall \alpha\, \exists \beta > \alpha\, \forall \overline{v} \in \mathbf{L}_\beta\, [\exists z \in \mathbf{L}\, \psi^{\mathbf{L}}(z, \overline{v}) \to \exists z \in \mathbf{L}_\beta\, \psi^{\mathbf{L}}(z, \overline{v})],$$

i.e.

$$\forall \alpha\, \exists \beta > \alpha\, \forall \overline{v} \in \mathbf{L}_\beta\, \exists z \in \mathbf{L}_\beta\, [\exists z \in \mathbf{L}\, \psi^{\mathbf{L}}(z, \overline{v}) \to \psi^{\mathbf{L}}(z, \overline{v})].$$

We will reason in ZF, taking care that all steps are applications of general ZF-theorems that do not depend on the specific formula ψ. Take any ordinal α. We know using only predicate logic that

$$\forall \overline{v} \in \mathbf{L}_\alpha\, \exists z \in \mathbf{L}\, [\exists z \in \mathbf{L}\, \psi^{\mathbf{L}}(z, \overline{v}) \to \psi^{\mathbf{L}}(z, \overline{v})];$$

therefore,

$$\forall \overline{v} \in \mathbf{L}_\alpha\, \exists! \alpha_{\overline{v}}\, (\alpha_{\overline{v}} = \bigcap \{\beta > \alpha \mid \exists z \in \mathbf{L}_\beta\, [\exists z \in \mathbf{L}\, \psi^{\mathbf{L}}(z, \overline{v}) \to \psi^{\mathbf{L}}(z, \overline{v})]\}).$$

by the unrelativized replacement and union axioms, there is a β_1 such that $\beta_1 = \sup\{\alpha_{\overline{v}} | \overline{v} \in \mathbf{L}_\alpha\}$. Continuing in this way, we can define by recursion a sequence β_p for $p \in \omega$, where for all $p \in \omega$,

$$\forall \overline{v} \in \mathbf{L}_{\beta_p}\, \exists z \in \mathbf{L}_{\beta_{p+1}}\, [\exists z \in \mathbf{L}\, \psi^{\mathbf{L}}(z, \overline{v}) \to \psi^{\mathbf{L}}(z, \overline{v})] \tag{2}$$

Define $\beta := \sup\{\beta_p \mid p \in \omega\}$. Because $\alpha = \beta_0 < \beta_1 < \beta_2 < \ldots$, we infer that β is a limit ordinal $> a$. Now using(2) and the fact that $\mathbf{L}_\beta = \bigcup_{\gamma < \beta} \mathbf{L}_\gamma$, we find that

$$\forall \overline{v} \in \mathbf{L}_\beta\, \exists z \in \mathbf{L}_\beta\, [\exists z \in \mathbf{L}\, \psi^{\mathbf{L}}(z, \overline{v}) \to \psi^{\mathbf{L}}(z, \overline{v})],$$

as desired. QED

Lemma 2.11 *For all φ, ZF feasibly proves the comprehension schema for φ, relativized to $\mathbf{L}$; i.e. by proofs of length polynomial in $|\varphi|$, ZF proves the following:*

$$\forall z, \overline{v} \in \mathbf{L}\, \exists y \in \mathbf{L}\, \forall x \in \mathbf{L}\, [x \in y \leftrightarrow x \in z \wedge \varphi^{\mathbf{L}}(x, z, \overline{v})]$$

Proof. Combine lemmas 2.5, 2.6 and 2.10. QED

Lemma 2.12 *For all φ, ZF feasibly proves the replacement schema for φ, relativized to $\mathbf{L}$; i.e. by proofs of length polynomial in $|\varphi|$, ZF proves the following:*

$$\begin{aligned}\forall a, \overline{v} \in \mathbf{L}\quad & [\forall x \in a\, \exists! y \in \mathbf{L}\, \varphi^{\mathbf{L}}(x, y, \overline{v}) \rightarrow \\ & \exists c \in \mathbf{L}\, \forall y \in \mathbf{L}\, (y \in c \leftrightarrow \exists x \in a\, \varphi^{\mathbf{L}}(x, y, \overline{v}))]\end{aligned}$$

Proof. We already have feasible proofs of the relativized comprehension schema for the formula $y \in b \wedge \exists x \in a\, \varphi(x, y, \overline{v})$. So we can (feasibly) prove that it suffices to show the following by proofs of length polynomial in $|\varphi|$:

$$ZF \vdash \forall a, \overline{v} \in \mathbf{L}\, [\forall x \in a\, \exists! y \in \mathbf{L}\, \varphi^{\mathbf{L}}(x, y, \overline{v}) \rightarrow \exists b \in \mathbf{L}\, (\forall x \in a\, \exists y \in b\, \varphi^{\mathbf{L}}(x, y, \overline{v}))]$$

The last proof works, as in lemma 2.10, by general theorems of ZF that do not depend on the specific φ. Work in ZF and suppose $a, \overline{v} \in \mathbf{L}$ and $\forall x \in a\, \exists! y \in \mathbf{L}\, \varphi^{\mathbf{L}}(x, y, \overline{v})$. Now

$$\forall x \in a\, \exists! \beta_x\, (\beta_x = \bigcap \{\alpha \mid \exists y \in \mathbf{L}_\alpha\, \varphi^{\mathbf{L}}(x, y, \overline{v})\});$$

then by replacement and the union axiom we find β such that $\beta = \bigcup \{\beta_x \mid x \in a\}$, and we let b be $\mathbf{L}_\beta$. Then

$$\forall x \in a\, \exists y \in b\, \varphi^{\mathbf{L}}(x, y, \overline{v}).$$

QED

Contrary to our expectations, the usual interpretation of $ZF + \mathbf{V} \neq \mathbf{L}$ into $ZF(M)$ (by forcing with generic extensions), although much more complex, is still feasible. We checked this following the lines of the proof in [Ku 80]. Our proof relies so heavily on the many details of Kunen's proof, that it would be incomprehensible to readers not conversant with that book. Therefore, we do not give it here.

In the literature there are also proofs of $ZF \rhd ZF + \mathbf{V} \neq \mathbf{L}$ and $ZF + AC \rhd ZF + AC + \neg CH$ which entirely avoid the use of the transitive countable collection M. A sketch of such a proof can be found in [Co 66, Section IV.11], and a completely different full proof in [VH 72, Ch. V, VI]. It appears that these proofs can also be analyzed to show that the interpretations in question are feasible.

Other well-known interpretations, such as the one of PA into ZF, are also feasible, as the reader may check for her/himself. All in all it seems that the only examples of theories U and V such that $U \rhd V$ but not $U \rhd_f V$ are contrived theories obtained by fixed-point constructions like the ones in section 4. It would be nice to find a more natural counterexample.

It would also be interesting to investigate severely restricted kinds of interpretability which do distinguish between interpretations used in everyday mathematics. For example, one could restrict the complexity of formulas allowed to occur in the proofs of the interpreted axioms.

Sam Buss suggested the following restricted definition of feasible interpretability to us:

$$U \rhd_{fm} V \leftrightarrow \exists K \exists M(\text{“}K \textit{ is an interpretation and } M \textit{ is a deterministic polynomial time Turing Machine”} \wedge \forall a(\alpha_V(a) \rightarrow Prf_U(M(a), a^K))). \qquad (3)$$

This definition is more in line with the conventional use of the word "feasible" in the context of polynomial time computability. The clause $Prf_U(M(a), a^K)$ in (3) is a P-like formula, while the clause $\exists p(\text{“}|p| \leq P(|a|)\text{”} \wedge Prf_U(p, a^K))$ in the definition of feasible interpretability used in this paper is an NP-like formula. However, all interpretations considered in this section can also be shown to be feasible in Buss's sense: we only need an easy analogue of lemma 2.2.

3 Soundness of ILM for feasible interpretability over PA

In this section, we restrict our attention to feasible interpretability over PA. We show that the modal interpretability logic ILM is PA-sound even if the intended meaning of $A \rhd B$ is "$PA + A$ feasibly interprets $PA + B$".

Definition 3.1 The modal interpretability logic ILM contains, besides all formulas having the form of a propositional tautology, the usual axioms for the provability logic L and the rules modus ponens and necessitation, the following axioms:

J1 $\Box(A \rightarrow B) \rightarrow (A \rhd B)$

J2 $(A \rhd B) \wedge (B \rhd C) \rightarrow (A \rhd C)$

J3 $(A \rhd C) \wedge (B \rhd C) \rightarrow (A \vee B \rhd C)$

J4 $(A \rhd B) \rightarrow (\Diamond A \rightarrow \Diamond B)$

J5 $\Diamond A \rhd A$

M $(A \rhd B) \rightarrow (A \wedge \Box C \rhd B \wedge \Box C)$

Definition 3.2 A *feasibility interpretation* is a map * which assigns to every propositional variable p a sentence p^* of the language of PA, and which is extended to all modal formulas as follows:

1. $(A \rhd B)^* = PA + A^* \rhd_f PA + B^*$
2. $(\Box A)^* = Prov_{PA}(A^*)$
3. * commutes with the propositional connectives.

Here $\rhd_f$ abbreviates the formalization of feasible interpretability

We will prove that ILM is arithmetically sound for feasible interpretability, i.e. that for all modal formulas A, if $ILM \vdash A$, then for all feasibility interpretions *, $PA \vdash A^*$. Thus, we have to check that the axioms J1 to J5 are valid in PA when $A \rhd B$ is read as $PA + A \rhd_f PA + B$. Whenever possible, we will prove generalizations of these axioms to theories $U, V \supseteq PA$. Also we prove a generalization of the property M, where an infinite set of Σ^0_1-sentences can be added on both sides instead of one $\Box$-sentence only.

Lemma 3.3 *PA proves all feasibility translations of J1 to J5.*

Proof. The proofs for J1 through J4 can be found almost verbatim in [Vi]. We reason in PA.

J1 Suppose for some theory V and some p that $Prf_V(p, \ulcorner A \urcorner)$. Then by the identity interpretation and the polynomial bound $P(n) = n + 3 \cdot |p|$, $V \rhd_f V + A$. So in particular, if $\Box_{PA}(A \to B)$, then $PA{+}A \rhd_f PA{+}A{+}B$, and surely $PA{+}A \rhd_f PA{+}B$.

J2 Suppose

- $U \rhd_f V$ by interpretation K_1 and polynomial P_1, and
- $V \rhd_f W$ by the interpretation K_2 and polynomial P_2.

As in the usual case, $U \rhd W$ by the interpretation $K_2 \circ K_1$. We need to show that there is a polynomial bound for the proofs of the translated axioms. So let b code an axiom of W, and p a proof in V of b^{K_2} with $|p| \le P_2(|b|)$.

If we take the K_1- translations of all formulas appearing in the proof coded by p, and add some intermediate steps, we can construct a U-proof of $(b^{K_2})^{K_1}$ from K_1-translations of axioms of V as assumptions; this proof will be of length $\le k \cdot |p|$, where k is a constant depending on the translation K_1 . Now we only have to add proofs of the translated V-axioms; the axioms themselves have codes of length $\le |p|$, so their K_1-translations have proofs with codes of length $\le P_1(|p|) \le P_1(P_2(|b|))$.

All in all, even in the worst case where the U-proof of $(b^{K_2})^{K_1}$ consists wholly of assumptions, there is a q with $|q| \le k \cdot P_2(|b|) \cdot P_1(P_2(|b|))$ such that $Prf_U(q, (b^{K_2})^{K_1})$. In particular, if $PA{+}A \rhd_f PA{+}B$ and $PA{+}B \rhd_f PA{+}C$, then $PA{+}A \rhd_f PA{+}C$.

J3 Suppose

- $U + A \rhd_f V$ by interpretation K_1 and polynomial P_1, and
- $U + B \rhd_f V$ by interpretation K_2 and polynomial P_2.

As in the usual case, we have $U{+}A{\vee}B \rhd V$ by the disjunctive interpretation M which equals K_1 in case A holds and equals K_2 in case $\neg A$ holds. To find a polynomial bound, we observe that for all C, $\vdash A \to (C^M \leftrightarrow C^{K_1})$ and $\vdash \neg A \to (C^M \leftrightarrow C^{K_1})$ by proofs of length $\le P(|C|)$, where the polynomial P depends on K_1 and K_2. Now suppose that c codes an axiom of V, that p_1 codes a $U + A$-proof of c^{K_1} with $|p_1| \le P_1(|c|)$, and that p_2 codes a $U + B$-proof of c^{K_2} with $|p_2| \le P_2(|c|)$. But then there is a constant k such that

- we can find p_1' such that $Prf_U(p_1', \ulcorner A \to \urcorner c^M)$ with $|p_1'| \le P(|c|) + P_1(|c|) + k \cdot |c|$; and
- we can find p_2' such that $Prf_U(p_2', \ulcorner \neg A \wedge B \to \urcorner c^M)$ with $|p_2'| \le P(|c|) + P_2(|c|) + k \cdot |c|$.

Combining p_1' and p_2' and their respective polynomial bounds, we find p and P' such that $Prf_U(p, \ulcorner A \vee B \to \urcorner c^M)$ with $|p| \le P'(|c|)$. In particular, we have: if $PA + A \rhd_f PA + C$ and $PA + B \rhd_f PA + C$, then $PA + A \vee B \rhd_f PA + C$.

J4 Because $(PA + A \rhd_f PA + B) \to (PA + A \rhd PA + B)$, we have by the soundness of J4 for normal interpretability immediately $(PA + A \rhd_f PA + B) \to (\Diamond A \to \Diamond B)$.

J5 In an easier variation of lemma 5.11, we use a claim proved in [Vi 89], which is stated in this paper as lemma 5.10. Suppose β is a Σ_1^b-formula axiomatizing a subset U of a Σ_1^b-language L. We will prove that $Q + \Diamond_\beta \top \rhd_f U$ i.e. $Q + \Diamond_U \top \rhd_f U$.

By lemma 5.10, we have

$$PA \vdash \Box_{Q+Con(\beta)} Con(\beta) \to \exists K \forall a \in Sent(L) Polprov_{Q+Con(\beta),|a|}(\ulcorner \Box_\beta a \to \urcorner a^K).$$

Of course we also know that $PA \vdash \Box_{Q+Con(\beta)} Con(\beta)$, so

$$PA \vdash \exists K \forall a \in Sent(L) Polprov_{Q+Con(\beta),|a|}(\ulcorner \Box_\beta a \to \urcorner a^K).$$

On the other hand, we have by provable Σ_1^b-completeness

$$PA \vdash \forall a(\beta(a) \to Polprov_{Q+Con(\beta),|a|}(\ulcorner \Box_\beta a \urcorner).$$

Combining the last two results, we have

$$PA \vdash \exists K \forall a(\beta(a) \to Polprov_{Q+Con(\beta),|a|}(a^K),$$

so $PA \vdash (Q + \Diamond_\beta \top) \rhd U$. In particular, we have for any sentence A:

$$PA \vdash (Q + \Diamond_{PA} A) \rhd_f PA + A,$$

so especially

$$PA \vdash PA + \Diamond_{PA} A \rhd_f PA + A.$$

QED

We want to prove that Montagna's property M holds for feasible interpretability over PA in its general version, where we can add an infinite set of Σ_1^0-sentences on both sides. In order to ensure that the usual arguments can indeed be polynomialized, we do not formulate the proof in the usual model-theoretic way, and we give many details that are not given in most proofs of Montagna's property for normal interpretability over PA. The example we give in theorem 4.1 of a set $\mathcal{S}$ of formulas such that $PA \vdash PA \rhd PA + \mathcal{S}$ but $\omega \not\models PA \rhd_f PA + \mathcal{S}$ also relies heavily on these details.

Suppose $U \supseteq PA$, $V \supseteq PA$. Now suppose $U \rhd_f V$ by the interpretation K (preserving $=$) with domain δ, and polynomial P. We want to find a polynomial Q such that for every Σ_1^0-sentence σ there is a $U + \sigma$-proof p of σ^K with $|\ulcorner p \urcorner| \leq Q(|\ulcorner \sigma \urcorner|)$. First, we need some definitions and lemmas. Fix U, V, K, P as given above.

Definition 3.4 Define *pism(s)* for "s is a partial isomorphism" and the function $G(j, y)$ as follows:

$$\begin{array}{rcl} pism(s) & := & seq(s) \wedge (s)_0 = 0^K \wedge \forall i < lh(s) - 1 (s)_{i+1} = S^K(s)_i \\ G(j,y) & := & \exists s(pism(s) \wedge lh(s) = j + 1 \wedge (s)_j = y) \end{array}$$

Lemma 3.5 *$U \vdash \forall j \exists! s(\mathrm{pism}(s) \wedge \mathrm{lh}(s) = j+1)$ and thus $U \vdash \forall j \exists! y G(j,y)$. Therefore, there is a function g corresponding to G.*

Proof. By induction. QED

Lemma 3.6 *U proves that g is injective, and $U \vdash \forall j \forall y(G(j,y) \rightarrow \delta(y))$.*

Proof. By induction. QED

Lemma 3.7 *U proves that g preserves $0, S, +, \cdot$, and $\leq$.*

Proof. We will give some of the preservation proofs. It follows immediately from the definition of *pism(s)* that $U \vdash g(0) = 0^K$ and $U \vdash \forall x(g(Sx) = S^K(g(x)))$.

We now prove by induction that g preserves $+$; the other proofs are analogous. We have $U \vdash g(x+0) = g(x) = g(x) +^K 0^K = g(x) +^K g(0)$ and $U \vdash g(x+y) = g(x) +^K g(y) \rightarrow g(x+Sy) = g(S(x+y)) = S^K(g(x+y)) = S^K(g(x) +^K g(y)) = g(x) +^K S^K(g(y)) = g(x) +^K g(Sy)$, So by induction (with x as parameter) $U \vdash \forall x \forall y(g(x+y) = g(x) +^K g(y))$. QED

Lemma 3.8 *The range of g is 'closed downwards', i.e. $U \vdash \forall x \forall u(\delta(u) \wedge u <^K g(x) \rightarrow \exists y < x(u = g(y)))$.*

Proof. Before we start the proof proper, we note a useful fact. V includes PA and K is an interpretation of V into U. Thus, as

1. $PA \vdash \forall x \forall u(u < x+1 \rightarrow u < x \vee u = x)$ and
2. $U \vdash \forall x(g(x) +^K 1^K = g(x+1))$, we also have
3. $U \vdash \forall x \forall u(\delta(u) \wedge u <^K g(x+1) \rightarrow u <^K g(x) \vee u = g(x))$.

Now we can start with the proof by induction on x of $U \vdash \forall x \forall u(\delta(u) \wedge u <^K g(x) \rightarrow \exists y < x(u = g(y)))$.

$\mathbf{x = 0}$ We have $U \vdash \neg \exists u(\delta(u) \wedge u <^K g(0))$, so $U \vdash \forall u(\delta(u) \wedge u <^K g(0) \rightarrow \exists y < 0(u = g(y)))$.

Induction step Work in U and suppose $\forall u(\delta(u) \wedge u <^K g(x) \rightarrow \exists y < x(u = g(y)))$ (induction hypothesis). Moreover, suppose $\delta(u) \wedge u <^K g(x+1)$. Then, by 3, $u <^K g(x) \vee u = g(x)$. So by the induction hypothesis $\exists y < x(u = g(y)) \vee u = g(x)$, i.e. $\exists y < x+1(u = g(y))$.

QED

Remark 3.9 Let $I(x)$ be the formula $\exists y(x = g(y))$. Note that U does not prove that I is closed under successor, so I does not define a cut; but by the previous lemma we do have $U \vdash \forall x \forall u(\delta(u) \wedge I(x) \wedge u <^K x \rightarrow I(u))$.

Lemma 3.10 *For all formulas $\varphi \in \Delta_1^0$, U proves the following by proofs of length polynomial in $|\ulcorner\varphi(a_1,\ldots,a_n)\urcorner|$:*

$$\varphi(a_1,\ldots,a_n) \leftrightarrow (\varphi^K)(g(a_1),\ldots,g(a_n)).$$

Proof. By induction on the construction of φ. We will see below that the proofs for the atomic formulae ψ are obviously of length linear in $|\ulcorner\psi\urcorner|$, and that all induction steps follow a given proof scheme in which the particular formulas at hand can be plugged in. So, because every φ has at most $|\ulcorner\varphi\urcorner|$ subformulas, there is a polynomial R such that for all φ, the U- proof of $\varphi(a_1,\ldots,a_n) \leftrightarrow (\varphi^K)(g(a_1),\ldots,g(a_n))$ is of length $\leq R(|\ulcorner\varphi\urcorner|)$. We will do the atomic step and the $\forall x \leq t$-step of the proof, and leave the others to the reader.

Atomic step By lemma 3.7, we have for all terms t by proofs of length polynomial in $|\ulcorner t\urcorner|$:

$$U \vdash g(t(a_1,\ldots,a_n)) = (t^K)(g(a_1),\ldots,g(a_n)).$$

So suppose φ is the formula $t_1(a_1,\ldots,a_n) = t_2(a_1,\ldots,a_n)$ where $a_1,\ldots,a_n$ include all variables appearing in t_1 and t_2. Then, because U proves that g is an injective function,

$$\begin{array}{lrcl} U \vdash & t_1(a_1,\ldots,a_n) & = & t_2(a_1,\ldots,a_n) \\ & \leftrightarrow g(t_1(a_1,\ldots,a_n)) & = & g(t_2(a_1,\ldots,a_n)) \\ & \leftrightarrow (t_1^K)(g(a_1),\ldots,g(a_n)) & = & (t_2^K)(g(a_1),\ldots,g(a_n)) \\ & \leftrightarrow ((t_1 & = & t_2)^K)(g(a_1),\ldots,g(a_n)) \end{array}$$

$\forall$x $\leq$ t-step Suppose that $\varphi(a_1,\ldots,a_n) = \forall x \leq t(a_1,\ldots,a_n)\psi(x,a_1,\ldots,a_n)$, and that $U \vdash \psi(x,a_1,\ldots,a_n) \leftrightarrow (\psi^K)(g(x),g(a_1),\ldots,g(a_n))$ (induction hypothesis). We will use the fact that, because of lemmas 3.6, 3.7 and 3.8,

$$\begin{array}{ll} U \vdash \forall u, a_1,\ldots,a_n & (\exists x\,[x \leq t(a_1,\ldots,a_n) \wedge u = g(x)] \\ & \leftrightarrow \delta(u) \wedge u \leq^K t^K(g(a_1),\ldots,g(a_n))). \end{array}$$

Thus, we have the following equivalences:

$$\begin{array}{ll} U \vdash & \varphi(a_1,\ldots,a_n) \\ & \leftrightarrow \forall x \leq t(a_1,\ldots,a_n)\psi(x,a_1,\ldots,a_n) \\ & \leftrightarrow \forall x \leq t(a_1,\ldots,a_n)(\psi^K)(g(x),g(a_1),\ldots,g(a_n)) \qquad \text{(by ind. hyp.)} \\ & \leftrightarrow \forall u(\delta(u) \wedge u \leq^K t^K(g(a_1),\ldots,g(a_n)) \rightarrow \psi^K(u,g(a_1),\ldots,g(a_n))) \\ & \leftrightarrow (\forall x \leq t\,\psi)^K(g(a_1),\ldots,g(a_n)) \qquad \text{(by def. of } K) \\ & \leftrightarrow (\varphi)^K(g(a_1),\ldots,g(a_n)) \end{array}$$

QED

Now we can finish the proof of the uniform version of Montagna's property M for feasible interpretability.

Theorem 3.11 *Suppose*

- U *satisfies full induction,*
- V *extends* PA *and*
- $U \rhd_f V$ *by interpretation* K *(preserving* $=$*) and polynomial* P.

Then there is a polynomial Q *such that for every* Σ_1^0*-sentence* σ *there is a* $U+\sigma$*-proof* p *of* σ^K *with* $|\ulcorner p \urcorner| \leq Q(|\ulcorner \sigma \urcorner|)$.

Thus, $U + \mathcal{S} \rhd_f V + \mathcal{S}$ *where* $\mathcal{S}$ *is a finite or infinite set of* Σ_1^0*-sentences.*

Proof. Suppose $\sigma \in \mathcal{S}$ is the Σ_1^0-sentence $\exists x \varphi(x)$, where $\varphi \in \Delta_1^0$. By lemma 3.10, there is a polynomial R such that we can prove the following by a proof of length $\leq R(|\ulcorner \sigma \urcorner|)$:

$$\begin{aligned} U \vdash \exists x \varphi(x) \ &\rightarrow \exists x \varphi^K(g(x)) \\ &\rightarrow \exists y(\delta(y) \wedge \varphi^K(y)) \\ &\rightarrow (\exists x \varphi(x))^K. \end{aligned}$$

Now we have $U + \mathcal{S} \rhd_f V + \mathcal{S}$ by the interpretation K and polynomial $Q := P + R$. QED

All results of this section also hold if we add the function symbol *exp* to the language of U and V, which we need in theorem 4.1. Let g be as defined in lemma 3.5. We will only give the result which needs some adaptation. The following preservation lemma corresponds to lemma 3.7:

Lemma 3.12 *Suppose* $exp \in L_U$. *Then* U *proves that* g *preserves* $0, S, +, \cdot, \leq$, *and* exp.

Proof. We already have a preservation proof for $\cdot$ by lemma 3.7. Preservation of *exp* then follows in the same way as preservation of $+$ was proved from preservation of S in lemma 3.7. QED

4 Interpretability does not imply feasible interpretability

Theorem 4.1 *There is a set* $\mathcal{S}$ *of* $\Delta_1^0(exp)$*-sentences such that* $\omega \models PA \rhd PA + \mathcal{S}$, *but* $\omega \not\models PA \rhd_f PA + \mathcal{S}$.

Proof. Define by Gödel's diagonalization theorem (or rather by the free variable version as formulated by Montague) a $\Delta_1^0(exp)$-formula $\varphi(y)$ such that

$$PA \vdash \varphi(y) \leftrightarrow \forall x \leq exp(y)\, \neg Prf(x, \ulcorner \varphi(\dot{y}) \urcorner).$$

It is easy to see that if we diagonalize directly, there is a polynomial O such that for each n, $|n| < |\ulcorner \varphi(\dot{n}) \urcorner| \leq O(|n|)$. Moreover, if $\varphi(\dot{n})$ were false, then by definition we would have a proof of the $\Delta_1^0(exp)$-sentence $\varphi(\dot{n})$; so $\varphi(\dot{n})$ must be true. But then, since $\varphi(\dot{n})$ is $\Delta_1^0(exp)$, we have the following:

1. PA proves $\varphi(\dot{n})$, though

2. because $\varphi(\dot{n})$ is true, PA does not prove $\varphi(\dot{n})$ by any proof whose Gödel number is of length $\leq n$.

Define $\mathcal{S} := \{\varphi(\dot{n}) \mid n \in \omega\}$. Then, by the identity interpretation, $\omega \models PA \rhd PA + \mathcal{S}$. Actually, as in [JM 88, section 6], we even have $PA \vdash \forall y Prov(\ulcorner\varphi(\dot{y})\urcorner)$, so $PA \vdash PA \rhd PA + \mathcal{S}$.

Now suppose, in order to derive a contradiction, that $\omega \models PA \rhd_f PA + \mathcal{S}$ by interpretation K and polynomial P. Thus, for all n,

$$PA \vdash \varphi(\dot{n})^K \text{by a proof of length} \leq P(|\ulcorner\varphi(\dot{n})\urcorner|).$$

We also know by lemma 3.10 (with $U = V = PA$) that there is a polynomial R such that for every n,

$$PA \vdash \varphi(\dot{n}) \leftrightarrow \varphi(\dot{n})^K \text{by a proof of length} \leq R(|\ulcorner\varphi(\dot{n})\urcorner|).$$

Now can construct from R and P a polynomial Q such that for all n,

$$PA \vdash \varphi(\dot{n}) \text{by a proof of length} \leq Q(|\ulcorner\varphi(\dot{n})\urcorner|).$$

However, there will be n such that $n > Q(O(|n|)) \geq Q(|\ulcorner\varphi(\dot{n})\urcorner|)$, and we have a contradiction with 2. QED

A salient feature of the counterexample above is the trivial identity interpretation by which PA interprets $PA + \mathcal{S}$. To prove that interpretability does not imply feasible interpretability, it is not essential that the set of formulas added to PA be infinite like $\mathcal{S}$ above. We will show a counterexample where one sentence can be normally but not feasibly interpreted over PA. Of course in this case the normal interpretation cannot be the identity. The counterexample also shows that in general we cannot feasibly merge two compatible feasible interpretations; i.e. it is not true that if $U \rhd_f V$, $U \rhd_f B$ and $U \rhd V + B$, then $U \rhd_f V + B$ (take $U = V = PA$, $B = A(\dot{n})$ or $B = E^*$ as below).

Theorem 4.2 *There is a sentence $A(\dot{n})$ such that $\omega \models PA \rhd PA + A(\dot{n})$, but $\omega \not\models PA \rhd_f PA + A(\dot{n})$.*

Proof. Let $P(x)$ be some Π^0_2-complete formula, say $P(x) = \forall y S(x, y)$, with $S \in \Sigma^0_1$. Define the formulas R and A by diagonalization such that

$$PA \vdash R(x, y) \leftrightarrow S(x, y) \preceq \Box_{PA} R(x, y)$$

and

$$PA \vdash A(x) \leftrightarrow \Box^*_{PA} \neg A(x) \preceq \exists y \neg R(x, y),$$

where $\Box^*$ is as defined in section 6.

Carrying out the proof of theorem 6.3 of the appendix section 6 in True Arithmetic, and taking the theory U mentioned there to be PA, we find the following result: if PA is consistent (as we believe it to be), then

$$\omega \models \forall x (PA \rhd PA + A(x) \leftrightarrow P(x)).$$

Now suppose, to derive a contradiction, that

$$\omega \models \forall x[(PA \rhd PA + A(x)) \leftrightarrow (PA \rhd_f PA + A(x))].$$

Then

$$\omega \models \forall x[(PA \rhd_f PA + A(x)) \leftrightarrow P(x)].$$

However, it is easy to see that $PA \rhd_f PA + A(x)$ is a Σ_2^0-predicate, contradicting the Π_2^0-completeness of P. Therefore, there is an $n \in \omega$ such that

- $\omega \models PA \rhd PA + A(\dot{n})$ but
- $\omega \not\models PA \rhd_f PA + A(\dot{n})$.

By this method we do not immediately find the value of a particular n that works, however.

A. Visser pointed out that we can make a specific counterexample in a more direct way using the Lindström method. Because $\{e \mid$ e is the Gödel number of a sentence E such that $\neg(PA \rhd_f PA + E)\} \in \Pi_2^0$, we can construct a formula A as in the appendix for which the following holds: for all sentences E,

$$\omega \models \neg(PA \rhd_f PA + E) \leftrightarrow PA \rhd PA + A(\ulcorner E \urcorner).$$

Now let E^* be the sentence constructed by the fixed point theorem such that

$$PA \vdash E^* \leftrightarrow A(\ulcorner E^* \urcorner).$$

Then

$$\omega \models \neg(PA \rhd_f PA + E^*) \leftrightarrow PA \rhd PA + E^*.$$

Therefore,

$$\omega \models PA \rhd PA + E^* \text{ and } \omega \not\models PA \rhd_f PA + E^*.$$

QED

5 ILM is the interpretability logic of feasible interpretability over PA

In this section, we will show that Berarducci's proof of the arithmetic completeness of ILM with respect to interpretability over PA can be adapted to prove that ILM is also arithmetically complete with respect to *feasible* interpretability over PA.

We have already proved in chapter 3 that for all modal formulas in the language of ILM we have:

if $ILM \vdash \varphi$, then for all feasibility interpretations *, $PA \vdash \varphi^*$.

Therefore, we will only need to show the converse:

if $ILM \nvdash \varphi$, then there is a feasibility interpretation * such that $PA \nvdash \varphi^*$.

We suppose that the reader has a copy of [Be 90] at hand in order to follow the original proofs. For the lemmas 5.5 up to 5.7, knowledge of [Pu 86], [Pu 87] or [Ve 89] will be helpful to the reader. As in [Pu 87], we take the logical complexity of a formula to be its quantifier depth. We can then adapt the results obtained in [Pu 87] to find for every standard n a formula Sat_n, a satisfaction predicate for formulas of logical complexity $\leq n$, such that Sat_n is of length linear in n. Subsequently, we can find proofs of length quadratic in n of the Tarski conditions and of the truth lemma for these satisfaction predicates Sat_n. Moreover, all these results can be formalized in PA. In the formalized case, we read Sat_n and $True_n$ as Gödel numbers found as function value in n. We will not go into the details here but refer the reader to the papers by Pudlák and Verbrugge.

First, we define some of the concepts that we use in the subsequent lemmas.

Definition 5.1 Formally, we define the following concepts:

- $Sent(a)$ for "a is the Gödel number of a sentence";
- $Fmla(a)$ for "a is the Gödel number of a formula";
- $Fmla_n(a)$ for "a is the Gödel number of a formula of logical complexity $\leq n$";
- $Cl(a)$ for "the Gödel number of the universal closure of the formula with Gödel number a"; note that Cl denotes a function;
- $Indax_n(b)$ for "b is the Gödel number of an induction axiom of logical complexity $\leq n$", i.e.

$$\begin{aligned} Indax_n(b) \iff\ & Fmla_n(b) \wedge \exists y[Fmla(y) \wedge \\ & b = Sub(y, \ulcorner v_1 \urcorner, \ulcorner 0 \urcorner) \ulcorner \wedge \forall v_1 (\urcorner y \ulcorner \rightarrow \urcorner Sub(y, \ulcorner v_1 \urcorner, \ulcorner S v_1 \urcorner) \ulcorner) \rightarrow \forall v_1 \urcorner y. \end{aligned}$$

We need to discriminate between a few different kinds of restricted provability, as defined below. In this section, provability means provability in PA, unless we explicitly state otherwise.

Definition 5.2 We formally define the following:

- $BPrf_n(x, y)$ for "x codes a proof of the formula coded by y, where only formulas of logical complexity $\leq n$ appear in the proof";
- $P\text{-}Polprov_n(x)$ for "x codes a formula that is provable by a proof of length $\leq P(n)$" where P is a polynomial;
- $Polprov_n(x)$ for "there is a polynomial P such that" $\forall n \exists p(|p| \leq P(n) \wedge Prf(p, x)$;
- $Polprov_{W,n}(x)$ for "there is a polynomial P such that "$\forall n \exists p(|p| \leq P(n) \wedge Prf_W(p, x)$
- $Prov_n(x)$ for "x codes a formula that is provable by a proof which only uses those axioms of PA with Gödel number $\leq n$"; abbreviation $\Box_n \varphi$ for $Prov_n(\ulcorner \varphi \urcorner)$;
- $Prov_{W,v}(x)$ for "x codes a formula that is provable by a proof which only uses those axioms of W with Gödel number $\leq n$"; abbreviation $\Box_{W,n} \varphi$ for $Prov_{W,n}(\ulcorner \varphi \urcorner)$.

In the context of satisfaction predicates $Sat_n(x, w)$, we need a few more concepts.

Definition 5.3 We formally define the following:

- $Evalseq(w, x)$ for "w encodes an evaluation sequence for the formula or term with Gödel number x; i.e. the length of the sequence w exceeds any i for which a variable v_i occurs in the formula or term coded by x";
- $s^*(i, x, w)$ for "the sequence which is identical to w, except that x appears in the i-th place"; note that s^* denotes a function;
- $True_n(x)$ for $\forall w(Evalseq(w, x) \to Sat_n(x, w))$;

Remark 5.4 When we prove formalized results, we read $True_n$ as a Gödel number just as Sat_n. So in that case the appropriate definition is as follows:

$True_n(x)$ for $\ulcorner \forall w(Evalseq(w, x) \to \urcorner Sat_n(x, w) \ulcorner) \urcorner$.

Lemma 5.5 (feasible subformula property) *There is a polynomial P such that*

$$PA \vdash \forall k \forall a(Fmla(a) \to P\text{-}Polprov_{|k|+|a|}[Prov_k(a) \to \exists q(BPrf_{|k|+|a|}(q, a))])$$

Proof. In [Ta 75], Takeuti gives a proof of the free cut-elimination theorem for PA, where PA is formulated as a Gentzen system. Free cut-elimination works in such a way that all principal formulas of induction inferences in the new free cut-free proof are substitution instances of principal formulas of induction inferences in the old proof. From this result Takeuti derives a proof of the corresponding subformula property for PA.

The proof of the subformula property can be adapted to the natural deduction formulation of PA, and can subsequently be formalized in PA. Thus, we can substitute any k bounding the Gödel numbers of axioms used, and any Gödel number a of a formula into the proof of the subformula property. Therefore, there is a polynomial P such that PA proves the following by proofs of length $\leq P(|k| + |a|)$:

$$PA \vdash Prov_k(a) \to \exists q(BPrf_{|k|+|a|}(q, a)).$$

Now this statement can again be formalized, so that we find

$$PA \vdash \forall k \forall a(Fmla(a) \to P\text{-}Polprov_{|k|+|a|}[Prov_k(a) \to \exists q(BPrf_{|k|+|a|}(q, a))]),$$

as desired. QED

Lemma 5.6 *There is a polynomial P such that*

$$PA \vdash \forall k \forall a(Fmla(a) \to P\text{-}Polprov_{|k|+|a|}[\exists q(BPrf_{|k|+|a|}(q, a) \to True_{|k|+|a|}(a)])$$

Proof. First, we work informally by induction on the construction of q. We work in PA, and we take any k and an a such that a is the Gödel number of a formula. We have to prove by polynomial length proofs (where the polynomial is fixed from outside) that $True_{|k|+|a|}$ preserves the axioms and rules as applied to formulas of logical complexity $\leq |k| + |a|$.

As an example, we show how this works for the induction schema. We take v_i as the induction variable in all our instances of the induction axioms. So suppose b codes an induction axiom of logical complexity $\leq |k| + |a|$, e.g. $b = (Sub(y, \ulcorner v_1 \urcorner, \ulcorner 0 \urcorner)^{\ulcorner} \wedge \forall v_1 (\urcorner y^{\ulcorner} \rightarrow \urcorner Sub(y, \ulcorner v_1 \urcorner, \ulcorner S v_1 \urcorner)^{\ulcorner}) \rightarrow \forall v_1 \urcorner y)$. We have to prove the following by a proof of length polynomial in $n := |k| + |a|$:

$$True_n(Sub(y, \ulcorner v_1 \urcorner, \ulcorner 0 \urcorner)^{\ulcorner} \wedge \forall v_1 (\urcorner y^{\ulcorner} \rightarrow \urcorner Sub(y, \ulcorner v_1 \urcorner, \ulcorner S v_1 \urcorner)^{\ulcorner}) \rightarrow \forall v_1 \urcorner y). \tag{4}$$

By a proof of length quadratic in n of the Tarski properties for Sat_n and a proof of length quadratic in n of a call by name / call by value lemma for Sat_n (cf. the proofs of lemmas 3.12 and 3.16 in [Ve 89],(4) is equivalent to the following:

$$\forall w[Sat_n(y, s^*(1,0,w)) \wedge \forall x(Sat_n(y, s^*(1,x,w)) \rightarrow Sat_n(y, s^*(1,Sx,w))) \rightarrow \\ \forall x(Sat_n(y, s^*(1,x,w))]. \tag{5}$$

The formulas (5) are themselves instances of induction of length linear in n, so they are provable by proofs of length linear in n. A polynomial of the form $P(n) = K \cdot n^3$ should now suffice to carry out the proofs of (4).

Again, we can formalize the argument to derive the following:

$$PA \vdash \forall k \forall a(Fmla(a) \rightarrow P\text{-}Polprov_{|k|+|a|}[\forall b(Indax_{|k|+|a|}(b) \rightarrow True_{|k|+|a|}(b))]).$$

Similarly, we can show by polynomially short proofs that the other axioms of logical complexity $\leq |k| + |a|$ are true, and that the rules preserve truth. We leave these proofs and their formalizations to the reader. QED

Lemma 5.7 *There is a polynomial P such that*

$$PA \vdash \forall k \forall a(Fmla(a) \rightarrow P\text{-}Polprov_{|k|+|a|}(True_{|k|+|a|}(a)^{\ulcorner} \rightarrow \urcorner Cl(a))$$

Proof. By a formalized Tarski's snowing lemma; cf. lemma 3.10 of [Ve 89]. QED

The following theorem corresponds to the reflection theorem 1.6 in [Be 90].

Theorem 5.8 (feasible reflection theorem) *There is a polynomial P such that*

$$PA \vdash \forall k \forall a(Sent(a) \rightarrow P\text{-}Polprov_{|k|+|a|}(\ulcorner Prov_k(a) \rightarrow \urcorner a))$$

Proof. Combine lemmas 5.5, 5.6 and 5.7. QED

In the following lemmas and theorems, $\exists K$ abbreviates $\exists K$(“K codes an interpretation” $\wedge \ldots$).

The next lemma was proved by Albert Visser [Vi 89, Chapter 6, Claim 3] in the course of a formalized Henkin construction in $I\Delta_0 + \Omega_1$.

Lemma 5.9 *Suppose β axiomatizes some subset of a Σ_1^b-language L. Then there is an r such that*

$$I\Delta_0 + \Omega_1 \vdash \Box_U Con(\beta) \rightarrow \exists K \forall a \in Sent(L) \exists p < \omega_1^r(a) Prf_U(p, \ulcorner \Box_\beta a \rightarrow \urcorner a^K).$$

Proof. See [Vi 89]. QED

Because of the correspondence between the values of ω_1-terms in a and exp(the values of polynomials in $|a|$), lemma 5.9 implies the following lemma:

Lemma 5.10 *Suppose β axiomatizes some subset of a Σ_1^b-language L. Then there is a polynomial P such that*

$$I\Delta_0 + \Omega_1 \vdash \Box_U Con(\beta) \rightarrow \exists K \forall a \in Sent(L) P\text{-}Polprov_{U,|a|}(\ulcorner \Box_\beta a \rightarrow \urcorner a^K).$$

The following theorem corresponds to Orey's theorem; see for example [Be 90, Theorem 2.9]

Theorem 5.11 (feasible Orey's theorem) *Suppose that $U \supseteq PA$ and W is axiomatized by α, where α is a Σ_1^b-formula. Then*

$$PA \vdash \forall x Polprov_{U,|x|}(\ulcorner \Diamond_{\alpha,x} \top \urcorner) \rightarrow U \rhd_f W.$$

Proof. Work in PA and suppose

$$\forall x Polprov_{U,|x|}(\ulcorner \Diamond_{\alpha,x} \top \urcorner).$$

In U, we will do a Henkin construction for the Feferman proof predicate for W. First define:

$$\beta(x) := \alpha(x) \wedge \Diamond_{\alpha,x+1} \top.$$

As in Feferman's original proof, we can prove that

$$\Box_U Con(\beta).$$

(For, reason in U and suppose $Prf_\beta(x, \perp)$, then for the axiom of β coded by the biggest Gödel number y to appear in x we have $\alpha(y) \wedge \neg \Diamond_{\alpha,y+1} \top$, thus $\neg \beta(y)$: a contradiction.)

On the other hand, by provable Σ_1^b-completeness for $\alpha(a)$ and by the assumption $\forall x Polprov_{U,|x|}(\ulcorner \Diamond_{\alpha,x} \top \urcorner)$, we have:

$$\forall a(\alpha(a) \rightarrow Polprov_{U,|a|}(\ulcorner \alpha(a) \wedge \Diamond_{\alpha,x+1} \top \urcorner)).$$

So, by definition of β, we have the following:

$$\forall a(\alpha(a) \rightarrow Polprov_{U,|a|}(\ulcorner \beta(a) \urcorner)). \tag{6}$$

But, using $\Box_U Con(\beta)$ we can apply lemma 5.10 to first derive

$$\exists K \forall a \in Sent(L) Polprov_{U,|a|}(\ulcorner \Box_\beta a \rightarrow \urcorner a^K),$$

and thus

$$\exists K \forall a \in Sent(L) Polprov_{U,|a|}(\ulcorner \beta(a) \rightarrow \urcorner a^K). \tag{7}$$

Finally we can combine 6 and 7 to get the desired conclusion

$$\exists K \forall a(\alpha(a) \rightarrow Polprov_{U,|a|}(a^K)),$$

i.e. $U \rhd_f W$. QED

Now we can start the proof of the arithmetical completeness of ILM with respect to feasible interpretations (cf. definition 3.2) over PA.

Theorem 5.12 *If $ILM \nvdash B$, then there is a feasibility interpretation * such that $PA \nvdash B^*$.*

The proof will in most places be identical to the one in [Be 90]. First we will sketch the outline of the proof, then we will prove the propositions that we need in the feasible case but differ essentially from those used in [Be 90].

Proof sketch. Suppose $ILM \nvdash B$, and take, by modal completeness of ILM with respect to simplified models, a provably primitive recursive ILM-Kripke model $V =< V, R, S, b, \Vdash>$, with $b = 1$ and $1 \Vdash B$. Extend V with a new root 0 with $0Rx$ for all $x \in V$, as in definition 5.1 of [Be 90]. Adapting definition 5.2 of [Be 90], we define a *feasibility interpretation* * such that for all propositional variables p,

$$p^* := \text{“}\exists x \in V \cup \{0\} : L = x \wedge x \Vdash p\text{”},$$

where L is defined as the limit of the Solovay function F, which is in turn defined in definition 5.7 of [Be 90]. We want to prove the following:

$$\text{whenever } 1 \nVdash A, \text{then } PA \nvdash A^*, \tag{8}$$

Then we will be done, as we have chosen V such that $1 \nVdash B$. To prove (8), we need to prove in PA a few properties of F and its limit L. Subsequently we need to prove by induction on the construction of the formula that for all formulas A, the feasibility interpretation * is *faithful* on A, i.e.

$$PA \vdash \forall x \in V(x \Vdash A \wedge L = x \rightarrow A^*) \text{ and}$$

$$PA \vdash \forall x \in V(x \Vdash \neg A \wedge L = x \rightarrow \neg A^*).$$

It is clear from the definition of F that * is faithful on atomic formulas. Moreover, the induction steps for the propositional connectives and $\Box$ immediately follow from the proofs in [Be 90]. Even the "negative" induction step for $\rhd$ has a straightforward proof:

> Work in PA and suppose $x \in V$, $x \Vdash \neg(A \rhd B)$, and $L = x$; then by part 2 in the proof of lemma 5.6 of [Be 90] and by the induction hypothesis, $\neg(A^* \rhd B^*)$. But then surely $\neg(A^* \rhd_f B^*)$, thus, as * is a feasibility interpretation, $\neg(A \rhd B)^*$.

For the "positive" direction, we need two extra lemmas. First we will prove in PA that F satisfies a feasible adaptation of Berarducci's property S , which we then use to finish the induction step for $\rhd$.

For $x \in V$, let $rank(x,n)$, the rank of x at stage n, be defined as in definition 5.7 of [Be 90]. The following proposition is an analogue of proposition 5.14 in [Be 90].

Proposition 5.13 (F has feasible property S) *PA proves the following:*

$$\begin{aligned} PA \vdash\ & \forall x \in V \cup \{0\}[L = x \rightarrow \\ & Polprov_{|k|}(\ulcorner \forall y, z \in V \cup \{0\}(L = y \wedge xRz \wedge ySz \rightarrow \Diamond_{\dot{k}} L = z)\urcorner)] \end{aligned}$$

Proof. We will prove the proposition by combining a few facts that are easy to check. For brevity's sake, we will leave out "$\in V \cup \{0\}$" after quantifiers $\forall x, \forall y, \forall z$. Likewise, capital P, with or without subscript, refers to formalized polynomials.

Fact 1 $PA \vdash Polprov_{|k|}(\ulcorner \forall y(L = y \rightarrow \Diamond_{\dot{k}} L = y)\urcorner)$

Proof. Immediately from the reflection theorem 5.8. The formula $L = y$ has a fixed length, so the polynomial found in the proof of the reflection theorem in this case depends only on $|k|$. QED

Fact 2 $PA \vdash Polprov_{|k|}(\ulcorner \forall y(L = y \rightarrow \forall n(\dot{k} < rank(y,n)))\urcorner)$

Proof. Immediately from fact 1 and the definition of rank. The appearance of k as an efficient numeral keeps the length of the proof polynomial in $|k|$. (This is also the case in the other facts below) QED

Fact 3 $PA \vdash Polprov_{|k|}(\ulcorner \forall z(\Box_{\dot{k}} L \neq z \rightarrow \exists m \forall n \geq m(rank(z,n) \leq \dot{k}))\urcorner)$

Proof. Immediately from the definition of rank. QED

Fact 4 $PA \vdash Polprov_{|k|}(\ulcorner \forall y \forall z(L = y \wedge \Box_{\dot{k}} L \neq z \rightarrow \exists n(F(n) = y \wedge n \textit{ codes } y \wedge rank(z,n) \leq \dot{k} \wedge rank(z,n) < rank(y,n)))\urcorner))$

Proof. From the definition of limit and fact 3: just take n big enough. We can take care that n codes y because we have an infinitely repetitive primitive recursive coding of the elements of $V \cup \{0\}$. Finally, to prove $rank(z,n) < rank(y,n)$, we use fact 2. QED

Fact 5 $PA \vdash \forall x(L = x \rightarrow Polprov_{|k|}(\ulcorner \exists j(J \geq \dot{k} \wedge F(j) = x)\urcorner))$

Proof. Immediate from the definition of the limit L of F. QED

Fact 6 We have the following:

$$\begin{aligned} PA \vdash\ & \forall x(L = x \rightarrow Polprov_{|k|}(\ulcorner \forall y \forall z(L = y \wedge \Box_{\dot{k}} L = z \wedge xRz \wedge ySz \rightarrow \\ & \exists n \exists j(F(n) = y \wedge n \textit{ codes } y \wedge rank(z,n) < rank(y,n) \wedge rank(z,n) \leq k \leq j \\ & \wedge F(j) = x \wedge F(rank(z,n))SxRz \wedge F(rank(z,n))Rz))\urcorner)) \end{aligned}$$

Proof. For the part up to $F(j) = x$, we combine facts 5 and 4. For the last two conjuncts, we use the monotonicity of F and the property corresponding to M of Veltman ILM-frames. QED

Fact 7 We have the following:

$$PA \vdash \forall x(L = x \rightarrow Polprov_{|k|} \ (\ulcorner \forall y \forall z(L = y \wedge \Box_{\dot{k}} L = z \wedge xRz \wedge ySz \rightarrow \exists n(F(n) = y \wedge F(n+1) = z))\urcorner)$$

Proof. Immediate from fact 6 and the definition of the function F, clause 2. QED

Now we can wrap up the proof: we see that $\exists n(F(n) = y \wedge F(n+1) = z)$ is inconsistent with $L = y$, so in fact we have what we were looking for:

$$PA \vdash \forall x[L = x \rightarrow Polprov_{|k|}(\ulcorner \forall y \forall z(L = y \wedge xRz \wedge ySz \rightarrow \Diamond_{\dot{k}} L = z)\urcorner)]$$

QED

The following proposition corresponds to part 1 of Lemma 5.6 of [Be 90].

Proposition 5.14 (positive induction step for $\rhd$) *Let * be the feasibility interpretation defined in the proof sketch of theorem 5.12. Suppose as induction hypothesis that*

$$PA \vdash \forall y(L = y \rightarrow (y \Vdash A \leftrightarrow A^*)) \text{ and}$$

$$PA \vdash \forall z(L = z \rightarrow (z \Vdash B \leftrightarrow B^*)).$$

Then

$$PA \vdash \forall x(L = x \wedge x \Vdash A \rhd B \rightarrow (A \rhd B)^*).$$

Proof. Let b be such that

$$PA \vdash \forall y(L = y \rightarrow (y \Vdash A \leftrightarrow A^*)) \text{ and}$$

$$PA \vdash \forall z(L = z \rightarrow (z \Vdash B \leftrightarrow B^*)),$$

both by proofs that use axioms of Gödel number up to b. Moreover suppose c is such that

$$PA \vdash \forall z(z \Vdash B \rightarrow \Box_c(z \Vdash B));$$

for this, any $c \geq$ the Gödel number of the biggest axiom of Robinson's arithmetic Q will do. Define $d := max(b, c)$. By theorem 5.11, the feasible version of Orey's theorem, it is sufficient to prove the following:

$$PA \vdash \forall x(L = x \wedge x \Vdash A \rhd B \rightarrow \forall k \geq d \ Polprov_{|k|}(\ulcorner A^* \rightarrow \Diamond_{\dot{k}} B^{*}\urcorner).$$

Again, we will state a list of easily provable facts from which the result immediately follows.

Fact 1 $PA \vdash \forall x(L = x \wedge x \Vdash A \rhd B \rightarrow \Box[A^* \rightarrow \exists y(L = y \wedge xRy \wedge y \Vdash A \wedge x \Vdash A \rhd B)])$

Proof. $L = x \rightarrow \Box\exists y(L = y \wedge xRy)$ by property $(\neg R)$, $\Box(A^* \wedge L = y \rightarrow y \Vdash A)$ by the induction hypothesis, and $\Box(x \Vdash A \rhd B)$ by provable Σ^0_1-completeness. QED

Fact 2 $PA \vdash \forall x(L = x \wedge x \Vdash A \rhd B \rightarrow \Box[A^* \rightarrow \exists y \exists z(L = y \wedge xRy \wedge y \Vdash A \wedge x \Vdash A \rhd B \wedge xRz \wedge ySz \wedge z \Vdash B)])$

Proof. From fact 1 and the definition of $x \Vdash A \rhd B$. QED

Fact 3 $PA \vdash \forall z \forall k \geq d\ Polprov_{|k|}(z \Vdash B \rightarrow \Box_{\dot{k}} z \Vdash B)$

Proof. From the assumption, and the fact that k appears only as efficient numeral. QED

Fact 4 $PA \vdash \forall x(L = x \wedge x \Vdash A \rhd B \rightarrow \forall k \geq d\ Polprov_{|k|}(\ulcorner A^* \rightarrow \exists y \exists z(L = y \wedge xRy \wedge xRz \wedge ySz \wedge \Diamond_{\dot{k}} L = z \wedge \Box_{\dot{k}} z \Vdash B)\urcorner))$

Proof. From fact 2 for $A^* \rightarrow \exists y \exists z(L = y \wedge xRy \wedge xRz \wedge ySz \wedge z \Vdash B)$; fact 3 for a proof of length polynomial in k of $z \Vdash B \rightarrow \Box_{\dot{k}} z \Vdash B)$, and proposition 5.13 for a proof of length polynomial in $|k|$ of $L = y \wedge xRy \wedge xRz \wedge ySz \rightarrow \Diamond_{\dot{k}} L = z$. QED

Fact 5 $PA \vdash \forall x(L = x \wedge x \Vdash A \rhd B \rightarrow \forall k \geq d\ Polprov_{|k|}(\ulcorner A^* \rightarrow \exists z \Diamond_{\dot{k}}(L = z \wedge z \Vdash B)\urcorner))$

Proof. If k is big enough (and $k \geq d$ will do), then by an easily formalized property of modus ponens, we have the following by proofs of length polynomial in $|k|$: $PA \vdash \forall z([\Box_k(z \Vdash B \rightarrow L \neq z) \wedge \Box_k z \Vdash B] \rightarrow \Box_k L \neq z)$, and thus $PA \vdash \forall z(\Diamond_k L = z \wedge \Box_k z \Vdash B \rightarrow \Diamond_k(L = z \wedge z \Vdash B))$. This argument can be formalized and combined with fact 4 to derive fact 5. QED

Fact 6 $PA \vdash \forall x(L = x \wedge x \Vdash A \rhd B \rightarrow \forall k \geq d\ Polprov_{|k|}(\ulcorner A^* \rightarrow \Diamond_{\dot{k}} B^{*}\urcorner)$

Proof. From fact 5 and the induction hypothesis; the fact that $k \geq d$ is used at this place. QED

From fact 5 and the feasible version of Orey's theorem, we may indeed derive

$$PA \vdash \forall x(x \Vdash A \rhd B \wedge L = x \rightarrow (A \rhd B)^*),$$

as desired.

QED

Proof sketch of theorem 5.12, continued. Concluding by induction that * is faithful on all formulas A, we have proved that $ILM \nvdash B^*$. Therefore, ILM is arithmetically complete with respect to feasible interpretability over PA.

QED.

References

[Be 90] A. Berarducci, The interpretability logic of Peano Arithmetic, *The Journal of Symbolic Logic*, vol. 55 (1990), pp. 1059-1089.

[Bu 86] S. Buss, *Bounded Arithmetic*, Bibliopolis, edizioni di filosofia e scienze, Napoli, 1986.

[Co 66] P.J. Cohen, *Set Theory and the Continuum Hypothesis*, Benjamin, New York, 1966.

[Fe 60] S. Feferman, Arithmetization of metamathematics in a general setting, *Fundamenta Mathematicae*, vol. 49 (1960), pp. 33-92.

[Há 79] P. Hájek, On partially conservative extensions of arithmetic, *Logic Colloquium 78*(M. Boffa et al., editors), North Holland, Amsterdam, 1979, pp. 225-234.

[JM 88] D. de Jongh and F. Montagna, Provable fixed points, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, vol. 34 (1988), pp. 229-250.

[JV 90] D. de Jongh and F. Veltman, Provability logics for relative interpretability, *Mathematical Logic* (*Proceedings, Chaika, Bulgaria, 1988*; P.P. Petkov, editor), Plenum Press, New York, 1990, pp. 31-42.

[Ku 80] K. Kunen, *Set Theory: An Introduction to Independence Proofs*, North-Holland, Amsterdam, 1980.

[Pu 85] P. Pudlák, Cuts, consistency statements and interpretability, *The Journal of Symbolic Logic*, vol. 50 (1985), pp. 423-441.

[Pu 86] P. Pudlák, On the length of proofs of finitistic consistency statements in first order theories, *Logic Colloquium '84* (J.B. Paris et al., editors), North Holland, Amsterdam, 1986, pp 165-196.

[Pu 87] P. Pudlák, Improved bounds on the length of finitistic consistency statements, *Logic and Combinatorics* (S.G. Simpson, editor), Contemporary Mathematics 35, American Mathematical Society, Providence, 1987, pp. 309-332.

[Sm 85] C. Smoryński, *Self-reference and Modal Logic*, Springer-Verlag, New York, 1985.

[So] R.M. Solovay, *On interpretability in set theories*, unpublished manuscript.

[Ta 75] G. Takeuti, *Proof Theory*, North Holland, Amsterdam, 1975.

[Ve 89] R. Verbrugge, Σ-completeness and bounded arithmetic, *ITLI Prepublication Series for Mathematical Logic and Foundations*, ML 89-05, University of Amsterdam, Amsterdam, 1989.

[Vi 88] A. Visser, Interpretability logic, *Logic Group Preprint Series*, nr. 40, University of Utrecht, Utrecht, 1988.

[Vi 89] A. Visser, The formalization of interpretability, *Logic Group Preprint Series*, nr. 47, University of Utrecht, Utrecht, 1989.

[Vi 90] A. Visser, *Proofs of Π_2-completeness by Per Lindström and Robert Solovay, as told by Albert Visser*, unpublished manuscript, 1990.

[Vi] A. Visser, *Questiones Longae et Breves*, unpublished cumulative manuscript.

[VH 72] P. Vopěnka and P. Hájek, *The Theory of Semisets*, North-Holland, Amsterdam, 1972.

6 Appendix

Solovay proved that the set $\{A \mid PA \rhd PA + A\}$ is Π^0_2-complete [So]. This result inspired Hájek to prove that, for every n, the set $\{A \mid A$ is Π^0_{n+1}-consevative over $PA\}$ is also Π^0_2-complete [Há 79].

We have adapted the proof of theorem 6.3 from Visser's unpublished rendition [Vi 90] of an alternative proof by Lindström of Hájek's general result.

Definition 6.1 Define $\Box_{U,x}B$ for "there is a proof of the formula B which only uses those axioms of U with Gödel number $\leq x$."

Suppose U and V are theories extending PA, such that for all B $PA \vdash \forall x \Box_U(\Box_{U,x}B \to B)$ (reflection for U), in particular $PA \vdash \forall x \Box_U \Diamond_{U,x}\top$. Then by the Orey-Hájek theorem, $PA \vdash U \rhd V \leftrightarrow \forall x \Box_U \Diamond_{V,x}\top$. The rest of the proof is taken almost verbatim from [Vi 90].

Let P be any Π_2-predicate, say $P = \forall x S(x)$, with $S \in \Sigma^0_1$. Pick R by diagonalization such that $PA \vdash R \leftrightarrow S \preceq \Box_U R$. Let $Q := \Box_U R \preceq S$. (We suppress free variables when convenient).

We first prove a lemma.

Lemma 6.2 $PA \vdash \Box_U R \leftrightarrow S \vee \Box_U \bot$.

Proof. Work inside PA and suppose $\Box_U R$. Then either R or Q holds. In case that R holds we have S by definition. In case that Q holds we have $\Box_U Q$ by Σ^0_1-completeness, and hence by definition both $\Box_U R$ and $\Box_U \neg R$, thus $\Box_U \bot$.

For the other direction, suppose S. Again we have either R or Q. From R we find $\Box_U R$ by Σ^0_1-completeness. From Q we immediately derive $\Box_U R$. Finally $\Box_U \bot$ gives $\Box_U R$ as well. QED

Define A by diagonalization such that $PA \vdash A \leftrightarrow \Box^*_U \neg A \preceq \exists y \neg R(y)$. Note that by lemma 6.2 we have $PA \vdash \Diamond_U \top \to [\forall x \Box_U R(x) \to P]$ and $PA \vdash P \to \forall x \Box_U R(x)$.

Theorem 6.3 $PA \vdash \Diamond_U \top \to (U \rhd U + A \leftrightarrow P)$

Proof. Work in PA and suppose $\Diamond_U \top$.

$\to$-side Suppose $U \rhd U + A$. Then by the Orey-Hájek theorem $\forall x \Box_U \Diamond_{U,x} A$. We will prove $\forall x \Box_U R(x)$. Pick any x. We have $\Box_U[Q(x) \to \neg R(x)]$; therefore by definition of A,

$$\Box_U[Q(x) \to \neg A \vee \Box_{U,x} \neg A]$$

and hence by reflection

$$\Box_U[Q(x) \to \neg A].$$

But then there is a y such that

$$\Box_{U,y}[Q(x) \to \neg A],$$

so by Σ^0_1- completeness

$$\Box_U \Box_{U,y}[Q(x) \to \neg A].$$

Also by Σ^0_1-completeness, there is a z such that

$$\Box_U[Q(x) \to \Box_{U,z} Q(x)].$$

Combining the previous two facts, we find a u such that

$$\Box_U[Q(x) \to \Box_{U,u} \neg A,]$$

and thus, by the assumption, $\Box_U \neg Q(x)$. It follows that

$$\Box_U[\Box_U R(x) \to R(x)],$$

hence by Löb's theorem $\Box_U R(x)$. We may conclude $\forall x \Box_U R(x)$, thus, because we have $\Diamond_U \top$, we conclude P.

$\leftarrow$-side Suppose P. Then $\forall x \Box_U R(x)$ and thus $\forall x \Box_U(\forall y < x R(y))$. It follows by definition of A that

$$\forall x \Box_U(\Box_{U,x} \neg A \to A).$$

On the other hand, we have

$$\forall x \Box_U(\Box_{U,x} \neg A \to \neg A)$$

by reflection, hence $\forall x \Box_U(\Diamond_{U,x} A)$. But then by the Orey-Hájek theorem $U \rhd U + A$.

QED

The ITLI Prepublication Series

LP-90-13 Zhisheng Huang — Logics for Belief Dependence
LP-90-14 Jeroen Groenendijk, Martin Stokhof — Two Theories of Dynamic Semantics
LP-90-15 Maarten de Rijke — The Modal Logic of Inequality
LP-90-16 Zhisheng Huang, Karen Kwast — Awareness, Negation and Logical Omniscience
LP-90-17 Paul Dekker — Existential Disclosure, Implicit Arguments in Dynamic Semantics

Mathematical Logic and Foundations

ML-90-01 Harold Schellinx — Isomorphisms and Non-Isomorphisms of Graph Models
ML-90-02 Jaap van Oosten — A Semantical Proof of De Jongh's Theorem
ML-90-03 Yde Venema — Relational Games
ML-90-04 Maarten de Rijke — Unary Interpretability Logic
ML-90-05 Domenico Zambella — Sequences with Simple Initial Segments
ML-90-06 Jaap van Oosten — Extension of Lifschitz' Realizability to Higher Order Arithmetic, and a Solution to a Problem of F. Richman
ML-90-07 Maarten de Rijke — A Note on the Interpretability Logic of Finitely Axiomatized Theories
ML-90-08 Harold Schellinx — Some Syntactical Observations on Linear Logic
ML-90-09 Dick de Jongh, Duccio Pianigiani — Solution of a Problem of David Guaspari
ML-90-10 Michiel van Lambalgen — Randomness in Set Theory
ML-90-11 Paul C. Gilmore — The Consistency of an Extended NaDSet

Computation and Complexity Theory

CT-90-01 John Tromp, Peter van Emde Boas — Associative Storage Modification Machines
CT-90-02 Sieger van Denneheuvel, Gerard R. Renardel de Lavalette — A Normal Form for PCSJ Expressions
CT-90-03 Ricard Gavaldà, Leen Torenvliet, Osamu Watanabe, José L. Balcázar — Generalized Kolmogorov Complexity in Relativized Separations
CT-90-04 Harry Buhrman, Edith Spaan, Leen Torenvliet — Bounded Reductions
CT-90-05 Sieger van Denneheuvel, Karen Kwast — Efficient Normalization of Database and Constraint Expressions
CT-90-06 Michiel Smid, Peter van Emde Boas — Dynamic Data Structures on Multiple Storage Media, a Tutorial
CT-90-07 Kees Doets — Greatest Fixed Points of Logic Programs
CT-90-08 Fred de Geus, Ernest Rotterdam, Sieger van Denneheuvel, Peter van Emde Boas — Physiological Modelling using RL
CT-90-09 Roel de Vrijer — Unique Normal Forms for Combinatory Logic with Parallel Conditional, a case study in conditional rewriting

Other Prepublications

X-90-01 A.S. Troelstra — Remarks on Intuitionism and the Philosophy of Mathematics, Revised Version
X-90-02 Maarten de Rijke — Some Chapters on Interpretability Logic
X-90-03 L.D. Beklemishev — On the Complexity of Arithmetical Interpretations of Modal Formulae
X-90-04 — Annual Report 1989
X-90-05 Valentin Shehtman — Derived Sets in Euclidean Spaces and Modal Logic
X-90-06 Valentin Goranko, Solomon Passy — Using the Universal Modality: Gains and Questions
X-90-07 V.Yu. Shavrukov — The Lindenbaum Fixed Point Algebra is Undecidable
X-90-08 L.D. Beklemishev — Provability Logics for Natural Turing Progressions of Arithmetical Theories
X-90-09 V.Yu. Shavrukov — On Rosser's Provability Predicate
X-90-10 Sieger van Denneheuvel, Peter van Emde Boas — An Overview of the Rule Language RL/1
X-90-11 Alessandra Carbone — Provable Fixed points in $I\Delta_0+\Omega_1$, revised version
X-90-12 Maarten de Rijke — Bi-Unary Interpretability Logic
X-90-13 K.N. Ignatiev — Dzhaparidze's Polymodal Logic: Arithmetical Completeness, Fixed Point Property, Craig's Property
X-90-14 L.A. Chagrova — Undecidable Problems in Correspondence Theory
X-90-15 A.S. Troelstra — Lectures on Linear Logic

1991

Logic, Semantics and Philosophy of Langauge

LP-91-01 Wiebe van der Hoek, Maarten de Rijke — Generalized Quantifiers and Modal Logic
LP-91-02 Frank Veltman — Defaults in Update Semantics
LP-91-03 Willem Groeneveld — Dynamic Semantics and Circular Propositions
LP-91-04 Makoto Kanazawa — The Lambek Calculus enriched with additional Connectives
LP-91-05 Zhisheng Huang, Peter van Emde Boas — The Schoenmakers Paradox: Its Solution in a Belief Dependence Framework
LP-91-06 Zhisheng Huang, Peter van Emde Boas — Belief Dependence, Revision and Persistence

Mathematical Logic and Foundations

ML-91-01 Yde Venema — Cylindric Modal Logic
ML-91-02 Alessandro Berarducci, Rineke Verbrugge — On the Metamathematics of Weak Theories
ML-91-03 Domenico Zambella — On the Proofs of Arithmetical Completeness for Interpretability Logic
ML-91-04 Raymond Hoofman, Harold Schellinx — Collapsing Graph Models by Preorders
ML-91-05 A.S. Troelstra — History of Constructivism in the Twentieth Century
ML-91-06 Inge Bethke — Finite Type Structures within Combinatory Algebras
ML-91-07 Yde Venema — Modal Derivation Rules
ML-91-08 Inge Bethke — Going Stable in Graph Models
ML-91-09 V.Yu. Shavrukov — A Note on the Diagonalizable Algebras of PA and ZF
ML-91-10 Maarten de Rijke, Yde Venema — Sahlqvist's Theorem for Boolean Algebras with Operators
ML-91-11 Rineke Verbrugge — Feasible Interpretability

Computation and Complexity Theory

CT-91-01 Ming Li, Paul M.B. Vitányi — Kolmogorov Complexity Arguments in Combinatorics
CT-91-02 Ming Li, John Tromp, Paul M.B. Vitányi — How to Share Concurrent Wait-Free Variables
CT-91-03 Ming Li, Paul M.B. Vitányi — Average Case Complexity under the Universal Distribution Equals Worst Case Complexity
CT-91-04 Sieger van Denneheuvel, Karen Kwast — Weak Equivalence
CT-91-05 Sieger van Denneheuvel, Karen Kwast — Weak Equivalence for Constraint Sets
CT-91-06 Edith Spaan — Census Techniques on Relativized Space Classes
CT-91-07 Karen L. Kwast — The Incomplete Database
CT-91-08 Kees Doets — Levationis Laus
CT-91-09 Ming Li, Paul M.B. Vitányi — Combinatorial Properties of Finite Sequences with high Kolmogorov Complexity
CT-91-10 John Tromp, Paul Vitányi — A Randomized Algorithm for Two-Process Wait-Free Test-and-Set
CT-91-11 Lane A. Hemachandra, Edith Spaan — Quasi-Injective Reductions

Computational Linguistics

CL-91-01 J.C. Scholtes — Kohonen Feature Maps in Natural Language Processing
CL-91-02 J.C. Scholtes — Neural Nets and their Relevance for Information Retrieval
CL-91-03 Hub Prüst, Remko Scha, Martin van den Berg — A Formal Discourse Grammar tackling Verb Phrase Anaphora

Other Prepublications

X-91-01 Alexander Chagrov, Michael Zakharyaschev — The Disjunction Property of Intermediate Propositional Logics
X-91-02 Alexander Chagrov, Michael Zakharyaschev — On the Undecidability of the Disjunction Property of Intermediate Propositional Logics
X-91-03 V. Yu. Shavrukov — Subalgebras of Diagonalizable Algebras of Theories containing Arithmetic
X-91-04 K.N. Ignatiev — Partial Conservativity and Modal Logics
X-91-05 Johan van Benthem — Temporal Logic
X-91-06 — Annual Report 1990
X-91-07 A.S. Troelstra — Lectures on Linear Logic, Errata and Supplement
X-91-08 Giorgie Dzhaparidze — Logic of Tolerance
X-91-09 L.D. Beklemishev — On Bimodal Provability Logics for Π_1-axiomatized Extensions of Arithmetical Theories
X-91-10 Michiel van Lambalgen — Independence, Randomness and the Axiom of Choice
X-91-11 Michael Zakharyaschev — Canonical Formulas for K4. Part I: Basic Results
X-91-12 Herman Hendriks — Flexibele Categoriale Syntaxis en Semantiek: de proefschriften van Frans Zwarts en Michael Moortgat
X-91-13 Max I. Kanovich — The Multiplicative Fragment of Linear Logic is NP-Complete
X-91-14 Max I. Kanovich — The Horn Fragment of Linear Logic is NP-Complete
X-91-15 V. Yu. Shavrukov — Subalgebras of Diagonalizable Algebras of Theories containing Arithmetic, revised version
X-91-16 V.G. Kanovei — Undecidable Hypotheses in Edward Nelson's Internal Set Theory
X-91-17 Michiel van Lambalgen — Independence, Randomness and the Axiom of Choice, Revised Version