Institute for Language, Logic and Information

UNARY INTERPRETABILITY LOGIC

Maarten de Rijke

ITLI Prepublication Series
for Mathematical Logic and Foundations ML-90-04

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

Logic, Semantics and Philosophy of Language:
LP-88-01 Michiel van Lambalgen 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

Mathematical Logic and Foundations:
ML-88-01 Jaap van Oosten 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

Computation and Complexity Theory:
CT-88-01 Ming Li, Paul M.B.Vitanyi 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

Other prepublications:
X-88-01 Marc Jumelet On Solovay's Completeness Theorem

1989

Logic, Semantics and Philosophy of Language:
LP-89-01 Johan van Benthem 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

Mathematical Logic and Foundations:
ML-89-01 Dick de Jongh, Albert Visser 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 $I\Delta_0+\Omega_1$

Computation and Complexity Theory:
CT-89-01 Michiel H.M. Smid 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

Other Prepublications:
X-89-01 Marianne Kalsbeek 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 SEE INSIDE BACK COVER

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

UNARY INTERPRETABILITY LOGIC

Maarten de Rijke
Department of Mathematics and Computer Science
University of Amsterdam

ITLI Prepublications
for Mathematical Logic and Foundations
ISSN 0924-2090

Received June 1990

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

Unary Interpretability Logic

Maarten de Rijke*

Department of Mathematics and Computer Science
University of Amsterdam

May 1990

Abstract

Let T be an arithmetical theory. We introduce a unary modal operator 'I' to be interpreted arithmetically as the unary interpretability predicate over T. We present complete axiomatizations of the (unary) interpretability principles underlying two important classes of theories. We also prove some basic modal results about these new axiomatizations.

1 Introduction

The language $\mathcal{L}(\Box)$ of propositional modal logic consists of a countable set of proposition letters p_0, p_1, ..., and connectives $\neg$, $\wedge$ and $\Box$. $\mathcal{L}(\Box, \rhd)$ is the language of (binary) interpretability logic, and extends $\mathcal{L}(\Box)$ with a binary operator '$\rhd$'. ('$A \rhd B$' is read: 'A interprets B'.) The *provability logic L* is propositional logic plus the axiom schemas $\Box(A \to B) \to (\Box A \to \Box B)$, $\Box A \to \Box\Box A$ and $\Box(\Box A \to A) \to \Box A$, and the rules Modus Ponens ($\vdash A$, $\vdash A \to B \Rightarrow \vdash B$) and Necessitation ($\vdash A \Rightarrow \vdash \Box A$). The *binary interpretability logic IL* is obtained from L by adding the axioms

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

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

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

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

($J5$) $\Diamond A \rhd A$,

where $\Diamond \equiv \neg\Box\neg$. IL is taken as the base system; extensions of IL with one or more of the following schemas have also been studied:

(F) $A \rhd \Diamond A \to \Box\neg A$

(W) $A \rhd B \to A \rhd (B \wedge \Box\neg A)$

(M_0) $A \rhd B \to (\Diamond A \wedge \Box C) \rhd (B \wedge \Box C)$

(P) $A \rhd B \to \Box(A \rhd B)$

(M) $A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C$.

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

We use ILX to denote the system $IL+X$, where X is the name of some axiom schema. $ILMP$ denotes the system $IL+M+P$ plus the additional axiom $A \rhd B \rightarrow A \wedge (C \rhd D) \rhd B \wedge (C \rhd D)$. Let ILS be one of the systems introduced above; the system ILS^{ω} has as axioms all theorems of ILS plus all instances of the schema of *reflection*: $\Box A \rightarrow A$. Its sole rule of inference is Modus Ponens.

Recall that an L-frame is a pair $\langle W, R\rangle$ with $R \subseteq W^2$ transitive and conversely well-founded, and that an L-model is given by an L-frame $\mathcal{F}$ together with a forcing relation $\Vdash$ that satisfies the usual clauses for $\neg$ and $\wedge$, while $u \Vdash A$ iff $\forall v\,(uRv \Rightarrow v \Vdash A)$. A *(Veltman-) frame* for IL is a triple $\langle W, R, S\rangle$, where $\langle W, R\rangle$ is an L-frame, and $S = \{ S_w : w \in W \}$ is a collection of binary relations on W satisfying

1. S_w is a relation on wR
2. S_w is reflexive and transitive
3. if w', $w'' \in wR$ and $w'Rw''$ then $w'S_w w''$.

An *IL-model* is given by a Veltman-frame $\mathcal{F}$ for IL together with a forcing relation $\Vdash$ that satisfies the above clauses for $\neg$, $\wedge$ and $\Box$, while

$$u \Vdash A \rhd B \Leftrightarrow \forall v\,(uRv \,\text{and}\, v \Vdash A \Rightarrow \exists w\,(vS_u w \,\text{and}\, w \Vdash B)).$$

An ILP-model is an IL-model that satisfies the extra condition: if $wRw'RuS_w v$ then $uS_{w'}v$. An ILM-model is an IL-model satisfying the extra condition: if $uS_w vRz$ then uRz. A model is an $ILMP$-model if it is both an ILM- and an ILP-model, while it also satisfies the condition: if $xRyS_x zRuS_y v$ then $uS_z v$.

In the sequel T denotes a theory which has a reasonable notion of natural numbers and finite sequences. The theories we consider are either Σ^0_1-sound essentially reflexive theories (like PA), or Σ^0_1-sound finitely axiomatized sequential theories (like GB).

An *arithmetical interpretation* $(\cdot)^T$ of $\mathcal{L}(\Box, \rhd)$ in the language of T is a map which assigns to every proposition letter p a sentence p^T in the language of T, and which is defined on other modal formulas as follows:

1. $(\bot)^T$ is '$0 = 1$';
2. $(\cdot)^T$ commutes with $\neg$ and $\wedge$;
3. $(\Box A)^T$ is a formalization of '$T \vdash (A)^T$';
4. $(A \rhd B)^T$ is a formalization of '$T + (A)^T$ interprets $T + (B)^T$'.

So the operator $\rhd$ is interpreted arithmetically as the *binary* interpretability predicate over T. Interpretability over T may also be studied as a *unary* predicate on finite extensions of T. Obviously, the modal analysis of the unary interpretability predicate in the spirit of Solovay's analysis of provability has to be undertaken using a *unary* modal operator. It was Craig Smoryński who first introduced an operator to be interpreted as the unary interpretability predicate. (The present investigations were inspired by questions of his.) Švejdar was subsequently the first one to introduce a binary operator to be interpreted as the binary interpretability relation.

It is clear that interpretability as a binary relation is the basic notion, since unary interpretability is reducible to it. On the modal side this leads to the following definition:

Definition 1.1 Define in $\mathcal{L}(\Box, \rhd)$ the unary interpretability operator '**I**' by $\mathbf{I}A := \top \rhd A$, and let $\mathcal{L}(\Box, \mathbf{I})$ extend $\mathcal{L}(\Box)$ with **I**.

So $x \Vdash \mathbf{I}A$ iff $\forall y\,(xRy \to \exists z\,(yS_x z \wedge z \Vdash A))$. And given a theory T, it follows from the definition of an arithmetical interpretation that $(\mathbf{I}A)^T$ is a formalization of '$T + (A)^T$ is interpretable in T'.

Definition 1.2 The *unary interpretability logic il* is obtained from the provability logic L by adding the axioms

$(I1)$ $\mathbf{I}\Box\bot$

$(I2)$ $\Box(A \to B) \to (\mathbf{I}A \to \mathbf{I}B)$

$(I3)$ $\mathbf{I}(A \vee \Diamond A) \to \mathbf{I}A$

$(I4)$ $\mathbf{I}A \wedge \Diamond\top \to \Diamond A$.

Several axioms have special names:

(f) $\mathbf{I}\Diamond\top \to \Box\bot$

(m) $\mathbf{I}A \to \mathbf{I}(A \wedge \Box\bot)$

(p) $\mathbf{I}A \to \Box\mathbf{I}A$.

We use *ilm* to denote the system $il + m$, and *ilp* to denote $il + p$. For other axiom schemas S we will simply refer to $ILS \cap \mathcal{L}(\Box, \mathbf{I})$ as *ils*. Let *ils* be one of the systems *il*, *ilm* or *ilp*. The system ils^ω has as axioms all theorems of *ils* plus all instances of the schema of *reflection*: $\Box A \to A$. Its sole rule of inference is Modus Ponens.

In Section 2 we prove that $il = IL \cap \mathcal{L}(\Box, \mathbf{I})$, $ilm = ILM \cap \mathcal{L}(\Box, \mathbf{I})$ and $ilp = ILP \cap \mathcal{L}(\Box, \mathbf{I})$—thereby establishing that *ilp* is the unary interpretability logic of all finitely axiomatized sequential theories that extend $\mathrm{I}\Delta_0 + \mathrm{SupExp}$, and that *ilm* is the unary interpretability logic of all essentially reflexive theories. It will turn out that *ilm* is in fact the unary interpretability logic of all 'reasonable' arithmetical theories. We end Section 2 with some remarks on the hierarchy of extensions of *il*.

Next, in Section 3 we study the closed fragment of $\mathcal{L}(\Box, \mathbf{I})$, and investigate the modalities in this language. We then state and prove Interpolation Theorems for *il*, *ilm* and *ilp*—from this we obtain Fixed Point Theorems for these logics in a standard way.

We end this section with two useful Propositions. Let *ils* be one of the systems *il*, *ilm* or *ilp*, and let ILS be the corresponding binary system. We first show that $ils \subseteq ILS \cap \mathcal{L}(\Box, \mathbf{I})$:

Proposition 1.3 *Let* $A \in \mathcal{L}(\Box, \mathbf{I})$. *If* $ils \vdash A$ *then* $ILS \vdash A$.

Proof. It suffices to show that for $S = \top$, P, M, we have $ILS \vdash ils$. We only show that $IL \vdash I1$ and that $ILM \vdash m$.

By $J1$, $J5$ and $J3$ we have

$$IL \vdash \Box\bot \vee \Diamond\Box\bot \rhd \Box\bot \tag{1}$$

Furthermore

$$\begin{aligned} IL \vdash \Box(\top \to (\top \wedge \Box\bot) \vee \Diamond(\top \wedge \Box\bot)) \quad &\Rightarrow \quad IL \vdash \Box(\top \to \Box\bot \vee \Diamond\Box\bot) \\ &\Rightarrow \quad IL \vdash \top \rhd \Box\bot \vee \Diamond\Box\bot, \text{ by } J1 \\ &\Rightarrow \quad IL \vdash \top \rhd \Box\bot, \text{ by } J2 \text{ and } (1). \end{aligned}$$

To prove that $ILM \vdash m$, we use the fact that in ILM we can derive $A \rhd B \rightarrow A \rhd B \wedge \Box\neg A$. (Cf. [9].) Therefore $ILM \vdash m$. QED.

Here are some theorems and a derived rule of the unary systems:

Proposition 1.4 *1. If $il \vdash A$ then $il \vdash \mathbf{I}A$. In particular, $il \vdash \mathbf{I}\top$.*

2. $il \vdash \Box A \rightarrow \mathbf{I}A$.

3. $il \vdash \mathbf{I}A \rightarrow \mathbf{I}(A \wedge \Box\neg A)$.

4. $il + f \subseteq ilm \subseteq ilp$.

Proof. Items 1, 2 and 3 are left to the reader. To prove item 4, note

$$\begin{array}{lll} ilp \vdash \mathbf{I}A & \rightarrow & \Box\mathbf{I}A \\ & \rightarrow & \Box(\Diamond\top \rightarrow \Diamond A), \text{ by } I4 \\ & \rightarrow & \Box(A \wedge \Box\neg A \rightarrow A \wedge \Box\bot) \\ & \rightarrow & \mathbf{I}(A \wedge \Box\neg A) \wedge \Box(A \wedge \Box\neg A \rightarrow A \wedge \Box\bot), \text{ by 3.} \\ & \rightarrow & \mathbf{I}(A \wedge \Box\bot), \text{ by } I2. \end{array}$$

That is, $ilp \vdash m$. This establishes the inclusion $ilm \subseteq ilp$. The inclusion $il + f \subseteq ilm$ is immediate. QED.

Assuming that il does indeed axiomatize $IL \cap \mathcal{L}(\Box, \mathbf{I})$, we find that $\vdash \mathbf{I}A \Rightarrow \vdash A$ is not a derived rule of il: we have $il \vdash \mathbf{I}\Box\bot$, but $il \nvdash \Box\bot$ because $IL \nvdash \Box\bot$.

2 Completeness

In this Section we prove il to be modally complete with respect to finite IL-models. We also prove modal and arithmetical completeness results for ilm and ilp. To prove the arithmetical completeness of ilm (ilp) we first show that ilm (ilp) is modally complete with respect to ILM- (ILP)-models; after that we appeal to the existing arithmetical completeness results for ILM (ILP).

2.1 Preliminaries

Our modal completeness proofs use *infinite* maximal consistent sets instead of the finite ones used, for example, to prove L or IL complete (in [6] and [2] respectively.) Our approach has the advantage that it can do without the large adequate sets employed there. In this subsection we establish some results that will provide us with the building blocks for constructing counter models in our modal completeness proofs.

We start with some definitions. For the remainder of this subsection let ils denote either il, ilm or ilp.

Definition 2.1 Let Γ, Δ be two maximal ils-consistent sets.

1. Δ is called a *successor* of Γ ($\Gamma \prec \Delta$) if
   (a) $A \in \Delta$ for each $\Box A \in \Gamma$
   (b) $\Box A \in \Delta$ for some $\Box A \notin \Gamma$
2. Δ is called a *C-critical successor* of Γ if
   (a) $\Gamma \prec \Delta$
   (b) $\mathbf{I}C \notin \Gamma$

(c) $\neg C, \Box\neg C \in \Delta$.

Note that successors of C-critical successors are C-critical successors as well. Moreover, any successor is a $\bot$-critical successor.

Definition 2.2 A set of formulas Φ is *adequate* if

1. if $B \in \Phi$, and C is a subformula of B, then $C \in \Phi$
2. if $B \in \Phi$, and B is no negation, then $\neg B \in \Phi$.

Let Φ be an adequate set. Then we say that a formula $\Diamond B$ is *almost in* Φ, if $\Diamond B \in \Phi$ or $\mathbf{I}B \in \Phi$ or $B \equiv \top$.

Proposition 2.3 *Let Γ be a maximal ils-consistent set such that $\Diamond C \in \Gamma$. Then there is a maximal ils-consistent successor Δ of Γ with $C, \Box\neg C \in \Delta$.*

Proof. Well-known (or cf. [6]). QED.

Proposition 2.4 *Let Γ be a maximal ils-consistent set with $\neg\mathbf{I}C \in \Gamma$. Then there is a maximal ils-consistent C-critical successor Δ of Γ with $\Box\bot \in \Delta$.*

Proof. Let Δ be a maximal consistent extension of

$$\{D : \Box D \in \Gamma\} \cup \{\neg C, \Box\neg C\} \cup \{\Box\bot\}.$$

Note that if such a Δ exists, it must be a C-critical successor of Γ: since

$$\{D : \Box D \in \Gamma\} \cup \{\Box\bot\} \subseteq \Delta$$

it is a successor of Γ; and because $\{\neg C, \Box\neg C\} \subseteq \Delta$ it is also C-critical.

We only have to prove $\{D : \Box D \in \Gamma\} \cup \{\neg C\} \cup \{\Box\bot\}$ consistent, since $\Box\bot$ implies $\Box\neg C$. Now, suppose that this set is inconsistent. Then there are $D_1, \ldots, D_m$ such that $D_1, \ldots, D_m, \neg C, \Box\bot \vdash \bot$. Then

$$\begin{aligned} D_1, \ldots, D_m \vdash \Box\bot \to C \quad &\Rightarrow \quad \Box D_1, \ldots, \Box D_m \vdash \Box(\Box\bot \to C) \\ &\Rightarrow \quad \Box D_1, \ldots, \Box D_m \vdash \mathbf{I}C, \text{ by } I1 \text{ and } I3. \end{aligned}$$

So $\Gamma \vdash \mathbf{I}C$. This contradicts the consistency of Γ. QED.

Proposition 2.5 *Assume that $\mathbf{I}C \in \Gamma$, and that Δ is a maximal ils-consistent E-critical successor of Γ. Then there is a maximal ils-consistent E-critical successor Δ' of Γ such that $C, \Box\neg C \in \Delta'$.*

Proof. Assume that there is no such Δ'. Then there are $\Box D_1, \ldots, \Box D_n \in \Gamma$ such that

$$D_1, \ldots, D_n, \neg E, \Box\neg E, C, \Box\neg C \vdash \bot,$$

so

$$\begin{aligned} D_1, \ldots, D_n &\vdash C \wedge \Box\neg C \to E \vee \Diamond E \\ \Box D_1, \ldots, \Box D_n &\vdash \Box(C \wedge \Box\neg C \to E \vee \Diamond E) \\ \Gamma &\vdash \Box(C \wedge \Box\neg C \to E \vee \Diamond E). \end{aligned} \tag{2}$$

Since $\mathbf{I}C \in \Gamma$, it follows from 1.4 that $\mathbf{I}(C \wedge \Box\neg C) \in \Gamma$. By (2) and $I2$ it follows that $\Gamma \vdash \mathbf{I}(E \vee \Diamond E)$, which, by $I3$, implies $\Gamma \vdash \mathbf{I}E$ and $\mathbf{I}E \in \Gamma$—but this contradicts the fact that $\mathbf{I}E \notin \Gamma$ by the existence of an E-critical successor of Γ. QED.

2.2 Modal completeness of *il*

Given some (infinite) maximal *il*-consistent set Γ and a finite adequate set Φ, we define the structure $\langle W_\Gamma, R\rangle$, which consists of pairs $\langle \Delta, \tau\rangle$. Here, the maximal consistent sets Δ are needed to handle the truth definition for formulas in $\Gamma \cap \Phi$. And the sequences of (pairs of) formulas τ are used to carefully index the pairs we add to W_Γ. In this way we make sure that $\langle W_\Gamma, R\rangle$ will be a finite tree.

For the time being, let Γ be an infinite maximal *il*-consistent set, and let Φ be a finite adequate set. We use $\bar{w}, \bar{v}, \ldots$ to denote pairs $\langle \Delta, \tau\rangle$. If $\bar{w} = \langle \Delta, \tau\rangle$, then $(\bar{w})_0 = \Delta$, $(\bar{w})_1 = \tau$. We write $\sigma \subseteq \tau$ for σ is an initial segment of τ, and $\sigma \subset \tau$ if σ is a proper initial segment of τ. Finally, $(\bar{w})_1{}^\frown(\bar{v})_1$ denotes the concatenation of $(\bar{w})_1$ and $(\bar{v})_1$.

Definition 2.6 Define W_Γ to be a minimal set of pairs $\langle \Delta, \tau\rangle$ such that

1. $\langle \Gamma, \langle\langle\rangle\rangle\rangle \in W_\Gamma$
2. if $\langle \Delta, \tau\rangle \in W_\Gamma$, $\Diamond B \in \Delta$ is almost in Φ and $C \in \Phi$, and if there is a maximal *il*-consistent C-critical successor Δ' of Δ with $B, \Box\neg B \in \Delta'$, then $\langle \Delta', \tau^\frown\langle\langle B, C\rangle\rangle\rangle \in W_\Gamma$ for *one* such Δ'.

Define R on W_Γ by putting $\bar{w}R\bar{v}$ iff $(\bar{w})_1 \subset (\bar{v})_1$. Define S on W_Γ by putting $\bar{v}S_{\bar{w}}\bar{u}$ iff for some B, B', C, τ and σ:

$$(\bar{v})_1 = (\bar{w})_1{}^\frown\langle\langle B, C\rangle\rangle^\frown\tau \text{ and } (\bar{u})_1 = (\bar{w})_1{}^\frown\langle\langle B', C\rangle\rangle^\frown\sigma.$$

Remark 2.7 In 2.6 the pairs $\langle B, C\rangle$ code the following: if $\langle \Delta', \tau^\frown\langle\langle B, C\rangle\rangle\rangle \in W_\Gamma$, then for some $\langle \Delta, \tau\rangle \in W_\Gamma$, Δ' is a C-critical successor of Δ, and $\langle \Delta', \tau^\frown\langle\langle B, C\rangle\rangle\rangle$ was added to W_Γ because $\Diamond B \in \Delta$ is almost in Φ.

Proposition 2.8 *1. W_Γ is finite.*

2. If $(\bar{w})_1 = (\bar{v})_1$ then $\bar{w} = \bar{v}$.

3. If $\bar{w}R\bar{v}$ then $(\bar{w})_0 \prec (\bar{v})_0$.

4. $\langle W_\Gamma, R\rangle$ is a tree.

5. $\langle W_\Gamma, R, S\rangle$ is an IL-frame.

6. If $\langle \Delta, \tau\rangle \in W_\Gamma$ and E occurs as the second component in some pair in τ, then $\neg E, \Box\neg E \in \Delta$.

Proof. 1. Since $|\Phi| = m$ for some finite m, it follows that for some finite n, $|\{\Diamond B \in \Gamma : \Diamond B \text{ is almost in } \Phi\}| = n$. So Γ gives rise to adding at most $n \cdot m$ new elements to W_Γ. Now each of these new elements contains at most $n-1$ formulas of the form $\Diamond B$, where $\Diamond B$ is almost in Φ. Hence, each such element will give rise to adding at most $(n-1)\cdot m$ new elements to W_Γ. Continuing in this way we see that $|W_\Gamma| \leq 1 + \Pi_{i=0}^{n-1}((n-i)\cdot m) < \omega$.

2. Induction on $\mathrm{lh}((\bar{w})_1) = \mathrm{lh}((\bar{v})_1)$.

3. Use item 2 to prove 3 with induction on $\max\big(\mathrm{lh}((\bar{w})_1), \mathrm{lh}((\bar{v})_1)\big)$.

4. To prove that $\langle W_\Gamma, R\rangle$ is a tree, note first that transitivity and asymmetry are straightforward, so we only prove that for each $\bar{w} \in W_\Gamma$ the set of its R-predecessors is finite and linear. Finitiness is immediate by item 1. To prove linearity, assume that $\bar{u}R\bar{w}$ and $\bar{v}R\bar{w}$. Then $(\bar{u})_1 \subset (\bar{w})_1$ and $(\bar{v})_1 \subset (\bar{w})_1$, so $(\bar{u})_1 \subseteq (\bar{v})_1$ or $(\bar{v})_1 \subseteq (\bar{u})_1$. If $(\bar{u})_1 = (\bar{v})_1$ then $\bar{u} = \bar{v}$ by item 2, and we are done. If $(\bar{u})_1 \neq (\bar{v})_1$ then either $(\bar{u})_1 \subset (\bar{v})_1$ or $(\bar{v})_1 \subset (\bar{u})_1$, that is: $\bar{u}R\bar{v}$ or $\bar{v}R\bar{u}$.

5. Left to the reader.

6. Induction on the construction of W_Γ. QED.

Theorem 2.9 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then $il \vdash A$ iff for all finite IL-models $\mathcal{M}$ we have $\mathcal{M} \models A$.*

Proof. Proving soundness is left to the reader. To prove completeness, assume that $il \nvdash A$. We want to produce an IL-model that refutes A. Let Φ be a finite adequate set containing $\neg A$, and let Γ be a maximal il-consistent set containing $\neg A$. Construct $\langle W_\Gamma, R, S\rangle$ as in 2.6. We complete the proof by putting $\bar{w} \Vdash p$ iff $p \in (\bar{w})_0$ and by proving that for all $F \in \Phi$ and $\bar{w} \in W_\Gamma$ we have $\bar{w} \Vdash F$ iff $F \in (\bar{w})_0$. The proof is by induction on F. We only consider the cases $F \equiv \Diamond B$ and $F \equiv \mathbf{I}C$.

If $F \equiv \Diamond B \in (\bar{w})_0$ we have to show that $\exists \bar{v}\,(\bar{w}R\bar{v} \wedge B \in (\bar{v})_0)$. Note first that $\Diamond B$ is almost in Φ, and that $\bot \in \Phi$. By 2.3 there is a successor Δ of $(\bar{w})_0$ with B, $\Box\neg B \in \Delta$. Moreover, Δ is a $\bot$-critical successor of $(\bar{w})_0$. For, $\Diamond B \in (\bar{w})_0$ implies $\Diamond\top \in (\bar{w})_0$, so $\mathbf{I}\bot \in (\bar{w})_0$ would imply $\Diamond\bot \in (\bar{w})_0$, by axiom $I4$—which is impossible; therefore, $\mathbf{I}\bot \notin (\bar{w})_0$. Furthermore, it is clear that $\neg\bot$, $\Box\neg\bot \in \Delta$. Put $\bar{v} := \langle \Delta, (\bar{w})_1{}^\frown\langle\langle B, \bot\rangle\rangle\rangle$. Then we may assume that $\bar{v} \in W_\Gamma$. It is clear that $\bar{w}R\bar{v}$ and $B \in (\bar{v})_0$ as required.

If $F \equiv \Diamond B \notin (\bar{w})_0$ then $\Box\neg B \in (\bar{w})_0$, and we have to show that $\forall \bar{v}\,(\bar{w}R\bar{v} \rightarrow \neg B \in (\bar{v})_0)$. But this is obvious from the definitions.

Assume $\mathbf{I}C \notin (\bar{w})_0$. Then $\neg\mathbf{I}C \in (\bar{w})_0$, and $\Diamond\top \in (\bar{w})_0$. By the induction hypothesis we have to show that $\exists \bar{v}\,(\bar{w}R\bar{v} \wedge \forall \bar{u}\,(\bar{v}S_{\bar{w}}\bar{u} \rightarrow \neg C \in (\bar{u})_0))$. Apply 2.4, with $\Gamma = (\bar{w})_0$, to obtain a C-critical successor Δ of $(\bar{w})_0$, and define $\bar{v} := \langle \Delta, (\bar{w})_1{}^\frown\langle\langle \top, C\rangle\rangle\rangle$. Since $\Diamond\top \in (\bar{w})_0$ is almost in Φ, we may assume that $\bar{v} \in W_\Gamma$. Furthermore, if $\bar{v}S_{\bar{w}}\bar{u}$ then C occurs as the second component in some pair in $(\bar{u})_1$, hence $\neg C \in (\bar{u})_0$, by 2.8.(5).

Assume $\mathbf{I}C \in (\bar{w})_0$. By the induction hypothesis we have to show that $\forall \bar{v}\,(\bar{w}R\bar{v} \rightarrow \exists \bar{u}\,(\bar{v}S_{\bar{w}}\bar{u} \wedge C \in (\bar{u})_0))$. So let $\bar{v} \in \bar{w}R$. Then $(\bar{v})_0 \succ (\bar{w})_0$ by 2.8.(3), so $\Diamond\top \in (\bar{w})_0$, and therefore $\Diamond C \in (\bar{w})_0$ by axiom $I4$. By construction $\bar{v}$ is E-critical for some $E \in \Phi$. Now, apply 2.5, with $\Gamma = (\bar{w})_0, \Delta = (\bar{v})_0$, to obtain an E-critical successor Δ' of $(\bar{w})_0$ that contains C, $\Box\neg C$. Since $\Diamond C$ is almost in Φ, we may assume that $\bar{u} = \langle \Delta', (\bar{w})_1{}^\frown\langle\langle C, E\rangle\rangle\rangle \in W_\Gamma$. Clearly, $\bar{u}$ does the job. QED.

Proposition 2.10 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then $IL \vdash A$ iff $il \vdash A$.*

Proof. By [2] we have for all $A \in \mathcal{L}(\Box, \rhd)$, $IL \vdash A$ iff for all finite IL-models $\mathcal{M}$, $\mathcal{M} \models A$. From this and 2.9 the Proposition follows. QED.

2.3 Modal and arithmetical completeness of *ilm*

To prove the modal completeness of *ilm* we need to adapt the construction used in proving *il* complete somewhat. The counter model we will construct in the completeness proof will consist of pairs $\langle \Delta, \tau\rangle$, where Δ is a maximal *ilm*-consistent set, and τ is a sequence of triples of formulas.

For the the time being we fix a maximal *ilm*-consistent set Γ and a finite adequate set Φ.

Definition 2.11 Define W_Γ to be a minimal set of pairs $\langle \Delta, \tau\rangle$ such that

1. $\langle \Gamma, \langle\langle\rangle\rangle\rangle \in W_\Gamma$.
2. If $\langle \Delta, \tau\rangle \in W_\Gamma$, $\Diamond B \in \Delta \cap (\Phi \cup \{\Diamond\top\})$, $C \in \Phi$ and if there exists a C-critical successor Δ' of Δ with B, $\Box\neg B \in \Delta'$, then for *one* such Δ', $\langle \Delta', \tau^\frown\langle\langle B, \bot, C\rangle\rangle\rangle \in W_\Gamma$.

3. If $\langle\Delta,\ \tau\rangle \in W_\Gamma$, $\mathbf{I}B \in \Delta\cap\Phi$, $C\in\Phi$ and if there exists a C-critical successor Δ' of Δ with $B,\ \Box\bot\in\Delta'$, then $\langle\Delta',\ \tau^\frown\langle\langle\bot,\ B,\ C\rangle\rangle\rangle \in W_\Gamma$, for *one* such Δ'.

Define R on W_Γ by putting $\bar{w}R\bar{v}$ if $(\bar{w})_1 \subset (\bar{v})_1$. Define S on W_Γ by putting $\bar{v}S_{\bar{w}}\bar{u}$ iff for some B, B', E, E', C, σ and σ'

$$(\bar{v})_1 = (\bar{w})_1{}^\frown\langle\langle B,\ E,\ C\rangle\rangle^\frown\sigma \text{ and } (\bar{u})_1 = (\bar{w})_1{}^\frown\langle\langle B',\ E',\ C\rangle\rangle^\frown\sigma'$$

and

$$\text{if } B\equiv\bot \text{ then } B'\equiv\bot,$$

and

$$\text{if } E'\equiv\bot \text{ then } B'\equiv B,\ E'\equiv E \text{ and } \sigma\subseteq\sigma'.$$

Remark 2.12 In 2.11 the triples $\langle B,\ E,\ C\rangle$ code the following: if $\langle\Delta',\ \tau^\frown\langle\langle B,\ E,\ C\rangle\rangle\rangle\in W_\Gamma$, then there is some $\langle\Delta,\ \tau\rangle\in W_\Gamma$ such that Δ' is a C-critical successor of Δ, and if $B\not\equiv\bot$ then $\langle\Delta',\ \tau^\frown\langle\langle B,\ E,\ C\rangle\rangle\rangle$ was added to W_Γ because $\Diamond B\in\Delta\cap(\Phi\cup\{\Diamond\top\})$; if $B\equiv\bot$ then $E\not\equiv\bot$ and $\langle\Delta',\ \tau^\frown\langle\langle B,\ E,\ C\rangle\rangle\rangle$ was added to W_Γ because $\mathbf{I}E\in\Delta\cap\Phi$.

Proposition 2.13 *1. W_Γ is finite.*

2. If $(\bar{v})_1 = (\bar{w})_1{}^\frown\langle\langle B,\ E,\ C\rangle\rangle^\frown\sigma$ then either $B\equiv\bot$ or $E\equiv\bot$ (but not both); and if $B\equiv\bot$ then $\Box\bot\in(\bar{v})_0$ and $\sigma=\langle\,\rangle$.

3. If $(\bar{w})_1 = (\bar{v})_1$ then $\bar{w}=\bar{v}$.

4. If $\bar{w}R\bar{v}$ then $(\bar{w})_0 \prec (\bar{v})_0$.

5. $\langle W_\Gamma,\ R,\ S\rangle$ is an ILM-frame.

6. If $\bar{v}=\langle\Delta,\ \tau\rangle\in W_\Gamma$ and C occurs as the third component in some triple in τ then $\neg C,\ \Box\neg C\in\Delta$.

Proof. Items 1, 2, 3, 4 and 6 are left to the reader. Let us check that $\langle W_\Gamma,\ R,\ S\rangle$ satisfies all the conditions to be an ILM-frame:

- it is easily seen that R is transitive and irreflexive—so by item 1 it is also conversely well-founded;
- $S_{\bar{w}}\subseteq \bar{w}R\times\bar{w}R$ is immediate;
- to show that $S_{\bar{w}}$ is reflexive and transitive, use item 2;
- to show that $\bar{w}R\bar{v}R\bar{u}$ implies $\bar{v}S_{\bar{w}}\bar{u}$, use item 2;
- finally, we have to show that $\bar{v}S_{\bar{w}}\bar{u}R\bar{z}$ implies $\bar{v}R\bar{z}$; so assume that $\bar{v}S_{\bar{w}}\bar{u}$. By definition there are B, B', B'', E, E', E'', C, C'', σ, σ' and σ'' such that

$$\begin{aligned}(\bar{v})_1 &= (\bar{w})_1{}^\frown\langle\langle B,\ E,\ C\rangle\rangle^\frown\sigma\\ (\bar{u})_1 &= (\bar{w})_1{}^\frown\langle\langle B',\ E',\ C\rangle\rangle^\frown\sigma'\\ (\bar{z})_1 &= (\bar{w})_1{}^\frown\langle\langle B',\ E',\ C\rangle\rangle^\frown\sigma'^\frown\langle\langle B'',\ E'',\ C''\rangle\rangle^\frown\sigma''\end{aligned}$$

  Obviously, $B'\not\equiv\bot$, for otherwise, by item 2, $\Box\bot\in(\bar{u})_0$, and, by item 4, $\bot\in(\bar{z})_0$. Therefore, by item 2, $E'\equiv\bot$—but then $B\equiv B'$, $E\equiv E'$ and $\sigma\subseteq\sigma'$. In other words: $(\bar{v})_1\subset(\bar{z})_1$, which means that $\bar{v}R\bar{z}$. QED.

Theorem 2.14 *Let $A\in\mathcal{L}(\Box,\mathbf{I})$. Then $ilm\vdash A$ iff for all finite ILM-models $\mathcal{M}$ we have $\mathcal{M}\models A$.*

Proof. As before we only prove completeness. Assume $ilm \nvdash A$. Let Φ be a finite adequate set that contains $\neg A$, and let Γ be a maximal *ilm*-consistent set with $\neg A \in \Gamma$. Construct $\langle W_\Gamma, R, S \rangle$ as in 2.11. Define a forcing relation $\Vdash$ on $\langle W_\Gamma, R, S \rangle$ by putting $\bar{w} \Vdash p$ iff $p \in (\bar{w})_0$. As before, we prove by induction on F that for all $F \in (\Phi \cup \{ \Diamond\top \})$ and $\bar{w} \in W_\Gamma$ we have $\bar{w} \Vdash F$ iff $F \in (\bar{w})_0$. We only consider the case $F \equiv \mathbf{I}B$. (The case $F \equiv \Diamond B$ is similar to the corresponding case in the proof of 2.9.)

Assume $F \equiv \mathbf{I}B \notin (\bar{w})_0$. By the induction hypothesis we have to show that $\exists \bar{v}\,(\bar{w}R\bar{v} \to \forall \bar{u}\,(\bar{v}S_{\bar{w}}\bar{u} \to \neg B \in (\bar{u})_0))$. Now $\mathbf{I}B \notin (\bar{w})_0$ implies $\Diamond\top \in (\bar{w})_0$. Moreover, by 2.4 there exists a B-critical successor Δ' of $(\bar{w})_0$. Since $\Diamond\top \in (\bar{w})_0 \cap (\Phi \cup \{ \Diamond\top \})$, we may assume that $\bar{v} \in W_\Gamma$, where

$$\bar{v} := \langle \Delta', (\bar{w})_1 {}^\frown \langle\langle \top, \bot, B \rangle\rangle \rangle.$$

Clearly, $\bar{w}R\bar{v}$. Finally, if $\bar{v}S_{\bar{w}}\bar{u}$ then $(\bar{u})_1 = (\bar{w})_1 {}^\frown \langle\langle B', E', B \rangle\rangle {}^\frown \sigma$ for some B', E' and σ. Therefore, by item 6 of 2.13, $\neg B \in (\bar{u})_0$, as required.

Assume that $F \equiv \mathbf{I}B \in (\bar{w})_0$. By the induction hypothesis we have to show that $\forall \bar{v}\,(\bar{w}R\bar{v} \to \exists \bar{u}\,(\bar{v}S_{\bar{w}}\bar{u} \wedge B \in (\bar{u})_0))$. So assume that $\bar{v} \in \bar{w}R$. Then $(\bar{v})_1 = (\bar{w})_1 {}^\frown \langle\langle B', E', C \rangle\rangle {}^\frown \sigma$ for some B', E', C and σ. By item 6 of 2.13, $(\bar{v})_0$ is C-critical. Now $\mathbf{I}B \in (\bar{w})_0$ implies $\mathbf{I}(B \wedge \Box\bot) \in (\bar{w})_0$, by axiom m. Apply 2.5 to find a C-critical successor Δ' of $(\bar{w})_0$ with $B, \Box\bot \in \Delta'$. Since $\mathbf{I}B \in (\bar{w})_0 \cap \Phi$, we may assume that $\bar{u} \in W_\Gamma$, where

$$\bar{u} := \langle \Delta', (\bar{w})_0 {}^\frown \langle\langle \bot, B, C \rangle\rangle \rangle.$$

Obviously, we have $\bar{v}S_{\bar{w}}\bar{u}$ and $B \in (\bar{u})_0$ as required. QED.

Proposition 2.15 *Let $A \in \mathcal{L}(\Box, \rhd)$. Then $ILM \vdash A$ iff $ilm \vdash A$.*

Proof. By [2] we have for all $A \in \mathcal{L}(\Box, \rhd)$, $ILM \vdash A$ iff for all finite ILM-models $\mathcal{M}$, $\mathcal{M} \models A$. From this and 2.14 the result follows. QED.

Theorem 2.16 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$, and let T be a Σ^0_1-sound essentially reflexive theory. Then $ilm \vdash A$ iff for all interpretations $(\cdot)^T$ of $\mathcal{L}(\Box, \mathbf{I})$ in the language of T, $T \vdash (A)^T$.*

Proof. By [1, Theorem 3.8] we have for all $A \in \mathcal{L}(\Box, \rhd)$, $ILM \vdash A$ iff for all interpretations $(\cdot)^T$ of $\mathcal{L}(\Box, \rhd)$ in the language of T, $T \vdash (A)^T$. From this and and 2.15 the result follows. QED.

Proposition 2.17 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then the following are equivalent:*

1. $ilm^\omega \vdash A$
2. $ILM^\omega \vdash A$
3. $ilm \vdash \left(\bigwedge_{\Box B \in \mathrm{Sub}(A)} (\Box B \to B) \wedge \bigwedge_{\mathbf{I}D \in \mathrm{Sub}(A)} \Diamond\top \right) \to A.$

Proof. The implication $1 \Rightarrow 2$ is trivial. By (the proof of) [1, Theorem 6.5] $ILM^\omega \vdash A$ implies

$$ILM \vdash (\bigwedge_{\Box B \in \mathrm{Sub}(A)} (\Box B \to B) \wedge \bigwedge_{C \rhd D \in \mathrm{Sub}(A)} (C \to \Diamond C)) \to A.$$

Since $A \in \mathcal{L}(\Box, \mathbf{I})$ this is equivalent to

$$ILM \vdash (\bigwedge_{\Box B \in \mathrm{Sub}(A)} (\Box B \to B) \wedge \bigwedge_{\mathbf{I}D \in \mathrm{Sub}(A)} \Diamond\top) \to A.$$

Together with 2.15 this yields the implication 2 $\Rightarrow$ 3. The implication 3 $\Rightarrow$ 1 is straightforward since $ilm^\omega \vdash \Box B \to B$ for all $B \in \mathcal{L}(\Box, \mathbf{I})$, so in particular $ilm^\omega \vdash \Box\bot \to \bot$, i.e., $ilm^\omega \vdash \Diamond\top$. QED.

Theorem 2.18 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$, and let T be a Σ^0_1-sound essentially reflexive theory. Then $ilm^\omega \vdash A$ iff for all interpretations $(\cdot)^T$ of $\mathcal{L}(\Box, \mathbf{I})$ in the language of T, $(A)^T$ is true in the standard model.*

Proof. By [1, Theorem 6.5] we have $ILM^\omega \vdash A$ iff for all interpretations $(\cdot)^T$ of $\mathcal{L}(\Box, \rhd)$ in the language of T, $(A)^T$ is true in the standard model. By 2.17 this implies the Theorem. QED.

Proposition 2.19 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then $ILW \vdash A$ iff $ilm \vdash A$.*

Proof. Since m is a substitution instance of the axiom W, the direction from right to left is immediate from 2.10. Conversely, if $ilm \nvdash A$, then $ILM \nvdash A$ by 2.15. Recall from the proof of 1.3 that $ILM \vdash W$, i.e. that $ILM \supseteq ILW$. It follows that $ILW \nvdash A$. QED.

Let us call an arithmetical theory a *reasonable* theory if it is Σ^0_1-sound, R^+_1-axiomatized and its natural numbers satisfy $I\Delta_0 + \Omega_1$. (Cf. [8] for details and motivation.)

Theorem 2.20 *The system ilm is the unary interpretability logic of all reasonable arithmetical theories.*

Proof. In [8, Section 6.2] it is shown that ILW is valid for arithmetic interpretations in all reasonable arithmetical theories, hence by 2.19, the same holds for *ilm*. Therefore, the unary interpretability logic of all reasonable arithmetics contains *ilm*. Since, by 2.16, *ilm* is the unary interpretability logic of PA, the converse inclusion holds as well. QED.

2.4 Modal and arithmetical completeness of *ilp*

In stead of proving *ilp* modally complete with respect to ILP-models we prove a stronger result, notably the modal completeness of *ilp* with respect to $ILMP$-models. The proof of this result is a slight variation on the modal completeness proof for *ilm*.

As before we fix a maximal *ilp*-consistent set Γ and a finite adequate set Φ.

Definition 2.21 Define W_Γ to be a minimal set of pairs $\langle \Delta, \tau \rangle$ such that

1. $\langle \Gamma, \langle\langle\rangle\rangle \rangle \in W_\Gamma$.
2. If $\langle \Delta, \tau \rangle \in W_\Gamma$, $\Diamond B \in \Delta \cap (\Phi \cup \{\Diamond\top\})$, $C \in \Phi$ and if there exists a C-critical successor Δ' of Δ with $B, \Box\neg B \in \Delta'$, then for *one* such Δ', $\langle \Delta', \tau^\frown\langle\langle B, \bot, C\rangle\rangle \rangle \in W_\Gamma$.
3. If $\langle \Delta, \tau \rangle \in W_\Gamma$, $\mathbf{I}B \in \Delta \cap \Phi$, $C \in \Phi$ and if there exists a C-critical successor Δ' of Δ with $B, \Box\bot \in \Delta'$, then $\langle \Delta', \tau^\frown\langle\langle \bot, B, C\rangle\rangle \rangle \in W_\Gamma$, for *one* such Δ'.

Define R on W_Γ by putting $\bar{w}R\bar{v}$ iff $(\bar{w})_0 \subset (\bar{v})_0$. Define S on W_Γ by putting $\bar{v}S_{\bar{w}}\bar{u}$ iff for some B, B', E, E', C, τ and σ

$$(\bar{v})_1 = (\bar{w})_1{}^\frown\tau^\frown\langle\langle B, E, C\rangle\rangle \text{ and } (\bar{u})_1 = (\bar{w})_1{}^\frown\tau^\frown\langle\langle B', E', C\rangle\rangle^\frown\sigma$$

and

$$\text{if } B \equiv \bot \text{ then } B' \equiv \bot$$

and

$$\text{if } E' \equiv \bot \text{ then } B' \equiv B,\ E' \equiv E.$$

Proposition 2.22 *1.* W_Γ *is finite.*

2. If $(\bar{v})_1 = (\bar{w})_1{}^\frown\tau^\frown\langle\langle B,\ E,\ C\rangle\rangle^\frown\sigma$ *then either* $B \equiv \bot$ *or* $E \equiv \bot$ *(but not both); and if* $B \equiv \bot$ *then* $\Box\bot \in (\bar{v})_0$ *and* $\sigma = \langle\,\rangle$.

3. If $(\bar{w})_1 = (\bar{v})_1$ *then* $\bar{w} = \bar{v}$.

4. If $\bar{w}R\bar{v}$ *then* $(\bar{w})_0 \prec (\bar{v})_0$.

5. $\langle W_\Gamma,\ R\rangle$ *is a tree.*

6. $\langle W_\Gamma,\ R,\ S\rangle$ *is an ILMP-frame.*

7. If $\bar{v} = \langle\Delta,\ \tau\rangle \in W_\Gamma$ *and C occurs as the third component in some triple in* τ, *then* $\neg C,\ \Box\neg C \in \Delta$.

Proof. We only prove item 6. The proof that $\langle W_\Gamma,\ R,\ S\rangle$ is an *ILM*-frame is similar to the proof of 2.13.(5); to prove that $\langle W_\Gamma,\ R,\ S\rangle$ is also an *ILP*-frame, we have to show that $\bar{w}R\bar{w}'R\bar{u}S_{\bar{w}}\bar{v}$ implies $\bar{u}S_{\bar{w}'}\bar{v}$—but this is immediate. So it remains to be proved that $xRyS_xzRuS_yv$ implies uS_zv. Reasoning as in 2.13.(5) we find that $xRyS_xzRu$ implies $xRy\underline{R}zRu$. Now, if $y = z$ then we trivially have uS_zv, and if yRz then we have uS_zv because $\langle W_\Gamma,\ R,\ S\rangle$ is an *ILP*-frame. QED.

Theorem 2.23 *Let* $A \in \mathcal{L}(\Box, \mathbf{I})$. *Then* $ilp \vdash A$ *iff for all finite ILMP-models* $\mathcal{M}$ *we have* $\mathcal{M} \models A$.

Proof. As before we only prove completeness. Assume that $ilp \nvdash A$. Let Φ be a finite adequate set that contains $\neg A$, and let Γ be a maximal *ilp*-consistent set with $\neg A \in \Gamma$. Construct $\langle W_\Gamma,\ R,\ S\rangle$ as in in 2.21. Define a forcing relation $\Vdash$ on $\langle W_\Gamma,\ R,\ S\rangle$ by putting $\bar{w} \Vdash p$ iff $p \in (\bar{w})_0$. As before, we prove by induction on F that for all $F \in \Phi \cup \{\Diamond\top\}$ and $\bar{w} \in W_\Gamma$ we have $\bar{w} \Vdash F$ iff $F \in (\bar{w})_0$. The case $F \equiv \Diamond B$ is similar to the corresponding case in the proof of 2.9. So we only consider the case $F \equiv \mathbf{I}B$.

The case that $F \equiv \mathbf{I}B \notin (\bar{w})_0$ is entirely analogous to the corresponding case in the proof of 2.14.

Assume that $F \equiv \mathbf{I}B \in (\bar{w})_0$. By the induction hypothesis we have to show that $\forall\bar{v}\,(\bar{w}R\bar{v} \to \exists\bar{u}\,(\bar{v}S_{\bar{w}}\bar{u} \wedge B \in (\bar{u})_0))$. So assume that $\bar{v} \in \bar{w}R$. Since $\langle W_\Gamma,\ R\rangle$ is a tree, we can find a unique immediate R-predecessor $\bar{w}'$ of $\bar{v}$. By axiom p we must have $\mathbf{I}B \in (\bar{w}')_0$, and so by axiom m, $\mathbf{I}(B \wedge \Box\bot) \in (\bar{w}')_0$. By construction there are $B',\ E',\ C' \in \Phi$ such that

$$(\bar{v})_1 = (\bar{w}')_1{}^\frown\langle\langle B',\ E'\ C'\rangle\rangle,$$

that is: $(\bar{v})_0$ is a C'-critical successor of $(\bar{w}')_0$. By 2.5 there exists a C'-critical successor Δ of $(\bar{w}')_0$ with $B,\ \Box\bot \in \Delta$. Since $\mathbf{I}B \in (\bar{w}')_0 \cap \Phi$, and $C' \in \Phi$ we may assume that $\bar{u} \in W_\Gamma$, where

$$\bar{u} := \big\langle\Delta,\ (\bar{w}')_1{}^\frown\langle\langle\bot,\ B,\ C'\rangle\rangle\big\rangle.$$

Obviously, we have $\bar{v}S_{\bar{w}}\bar{u}$ and $B \in (\bar{u})_0$ as required. QED.

Proposition 2.24 *Let* $A \in \mathcal{L}(\Box, \mathbf{I})$. *Then* $ILMP \vdash A$ *iff* $ilp \vdash A$.

Proof. If $ilp \vdash A$ then by 1.3 $ILP \vdash A$, and hence $ILMP \vdash A$. Conversely, if $ILMP \vdash A$, then for all (finite) *ILMP*-models $\mathcal{M}$, $\mathcal{M} \models A$. So by 2.23, $ilp \vdash A$. QED.

Proposition 2.25 *Let* $A \in \mathcal{L}(\Box, \mathbf{I})$. *Then* $ILP \vdash A$ *iff* $ilp \vdash A$.

Proof. The direction from right to left follows from 1.3. To prove the other direction, note that $ilp \nvdash A$ implies $ILMP \nvdash A$, by the previous Proposition, and this in turn implies $ILP \nvdash A$. QED.

Theorem 2.26 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$ and let T be a Σ_1^0-sound finitely axiomatized sequential theory that extends* $\mathrm{I}\Delta_0 + \mathrm{SupExp}$. *Then $ilp \vdash A$ iff for all interpretations $(\cdot)^T$ of $\mathcal{L}(\Box, \mathbf{I})$ in the language of T, $T \vdash (A)^T$.*

Proof. By 2.25 we have $ilp \vdash A$ iff $ILP \vdash A$, for all $A \in \mathcal{L}(\Box, \mathbf{I})$. By [9, Theorem 8.2] this is equivalent to: for all interpretations $(\cdot)^T$ of $\mathcal{L}(\Box, \rhd)$ in the language of T, $T \vdash (A)^T$. This implies the Theorem. QED.

Proposition 2.27 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then the following are equivalent:*

1. $ilp^\omega \vdash A$
2. $ILP^\omega \vdash A$
3. $ilp \vdash \left(\bigwedge_{\Box B \in \mathrm{Sub}(A)}(\Box B \to B) \wedge \bigwedge_{\mathbf{I}D \in \mathrm{Sub}(A)} \Diamond\top\right) \to A.$

Proof. The implication 1 $\Rightarrow$ 2 is trivial. By [4, Proposition 1.8.(1)] $ILP^\omega \vdash A$ implies

$$ILP \vdash \Big(\bigwedge_{\Box B \in \mathrm{Sub}(A)} (\Box B \to B) \wedge \bigwedge_{C \rhd D \in \mathrm{Sub}(A)} (C \to \Diamond C)\Big) \to A.$$

Since $A \in \mathcal{L}(\Box, \mathbf{I})$ this is equivalent to

$$ILP \vdash \Big(\bigwedge_{\Box B \in \mathrm{Sub}(A)} (\Box B \to B) \wedge \bigwedge_{\mathbf{I}D \in \mathrm{Sub}(A)} \Diamond\top\Big) \to A.$$

Together with 2.25 this yields the implication 2 $\Rightarrow$ 3. The implication 3 $\Rightarrow$ 1 is straightforward since $ilp^\omega \vdash \Box B \to B$ for all $B \in \mathcal{L}(\Box, \mathbf{I})$. QED.

Theorem 2.28 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$, and let T be a Δ_2-sound finitely axiomatized sequential theory that extends* $\mathrm{I}\Delta_0 + \mathrm{SupExp}$. *Then $ilp^\omega \vdash A$ iff for all interpretations $(\cdot)^T$ of $\mathcal{L}(\Box, \mathbf{I})$ in the language of T, $(A)^T$ is true in the standard model.*

Proof. By [4, Theorem 3.2] we have $ILP^\omega \vdash A$ iff for all interpretations $(\cdot)^T$ of $\mathcal{L}(\Box, \rhd)$ in the language of T, $(A)^T$ is true in the standard model. By 2.27 this yields the Theorem. QED.

2.5 On the hierarchy of extensions of *il*

In [2], [8] and [9] the following extensions of IL in $\mathcal{L}(\Box, \rhd)$ are considered:

$$IL \subset ILF \subset ILW \subset ILWM_0 \subset \begin{matrix} ILP \\ \\ ILM \end{matrix} \subset ILMP$$

(All inclusions are proper.)

As a corollary to 2.19 and 2.24 we find that this hierarchy partly collapses when we only consider formulas $A \in \mathcal{L}(\Box, \mathbf{I})$:

$$il \subset ilf \subset ilw = ilwm_0 = ilm \subset ilp = ilmp.$$

(Recall that $ilx = ILX \cap \mathcal{L}(\Box, \mathbf{I})$.)

To see that there is no total collapse we prove the following result:

Proposition 2.29 *1. $ilm \neq ilp$*
2. $ilf \neq ilm$
3. $il \neq ilf$

Proof. 1. It suffices to show that $ilm \nvdash \mathbf{I}A \to \Box\mathbf{I}A$. Consider Figure 1 below. Note first that the model is an ILM-model, but not an ILP-model. We clearly have $w \Vdash \mathbf{I}p$. However, b does not force $\mathbf{I}P$, for it has an R-successor (notably a) that is not S_b-succeeded by a point at which p holds—so $w \nVdash \Box\mathbf{I}p$.

2. To prove $ilf \neq ilm$ we use a construction due to Švejdar. (Cf. [7].) It suffices to show that $ILF \nvdash m$. Consider Figure 2. We claim that $w \Vdash F$, i.e., that $w \Vdash A \rhd \Diamond A \to \Box\neg A$, for all $A \in \mathcal{L}(\Box, \mathbf{I})$. Suppose that $w \Vdash A \rhd \Diamond A$. Then

(a) if $b \Vdash A$ then $a \Vdash A$

(b) $d \nVdash A$—otherwise $d \Vdash \Diamond A$, which is impossible

(c) for each B, $a \Vdash B$ iff $c \Vdash B$

(d) $c \nVdash A$—otherwise $c \Vdash \Diamond A$, which is impossible

(e) $a \nVdash A$, by (c) and (d)

(f) $b \nVdash A$, by (a) and (e)

(g) $w \Vdash \Box\neg A$, by (b), (d), (e) and (f).

On the other hand, $w \nVdash \mathbf{I}A \to \mathbf{I}(A \wedge \Box\bot)$, for we have $w \Vdash \mathbf{I}p$ while $w \nVdash \mathbf{I}(p \wedge \Box\bot)$, since b has no S_w-successor at which $p \wedge \Box\bot$ holds.

3. We have $ilf \vdash f$, i.e., $ilf \vdash \mathbf{I}\Diamond\top \to \Box\bot$, but $il \nvdash f$, as is clear from the model in Figure 3. QED.

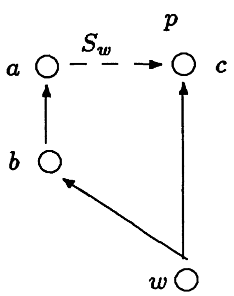


Figure 1.

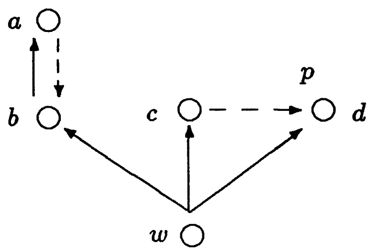


Figure 2.

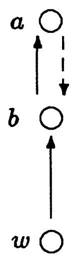


Figure 3.

(Plain arrows denote R-links; dashed arrows denote S_w-links; reflexive S-links and S-links induced by R-links have been left out.)

3 Answers to some standard questions

In this Section we answer some questions that come naturely with any extension of L. Notably, what are the closed formulas and the modalities in $\mathcal{L}(\mathbf{I})$ and $\mathcal{L}(\Box, \mathbf{I})$? We also prove interpolation and fixed point theorems for il, ilm and ilp.

3.1 Closed formulas and modalities

As usual we start with some definitions. A formula C is called *closed* if it does not contain any proposition letters. Let $\mathcal{F}$ be a frame. Define the *depth* $\mathrm{d}(w)$ of $w \in \mathcal{F}$ by $\mathrm{d}(w) = \sup\{\, \mathrm{d}(v) + 1 : wRv \,\}$.

Proposition 3.1 *Let w, v be two points (not necessarily in the same model). If $\mathrm{d}(w) = \mathrm{d}(v)$ then $w \Vdash C$ iff $v \Vdash C$ for all closed formulas $C \in \mathcal{L}(\Box)$.*

Proof. This is by induction on $\mathrm{d}(w) = \mathrm{d}(v)$. QED.

Proposition 3.2 *Let C be a closed formula in $\mathcal{L}(\Box)$. Then $L \vdash C$ iff C is valid on ω^*. (I.e., iff for every $w \in \omega^*$ and every $\Vdash$ on ω^*, $w \Vdash C$.)*

Proof. The direction from left to right is obvious. To prove the other one, assume that $L \nvdash C$; then for some finite L-model $\mathcal{M}$ with root w, $w \nVdash C$. Let $n = \mathrm{d}(w)$, and let $\Vdash'$ be any forcing relation on ω^*. It is clear that, in ω^*, the element n has depth n. So by the previous Proposition, $n \nVdash' C$. QED.

Proposition 3.3 *Let C be a closed formula in $\mathcal{L}(\Box)$. Then $L \vdash (C \vee \Diamond C) \leftrightarrow \Diamond^k \top$, for some $k \in \omega \cup \{\,\omega\,\}$. (Here, $\Diamond^\omega \top \equiv \bot$.)*

Proof. By the previous Proposition it suffices to show that for all closed formulas C in $\mathcal{L}(\Box)$, there is some $k \in \omega \cup \{\,\omega\,\}$ such that $(C \vee \Diamond C) \leftrightarrow \Diamond^k \top$ is valid on ω^*. This is left to the reader. QED.

Proposition 3.4 *Let X be a logic that extends $il+f$. Then every closed formula in $\mathcal{L}(\mathbf{I})$ is, provably in X, equivalent to one of $\Diamond\top$, $\Box\bot$, $\bot$ or $\top$. Hence, every closed formula in $\mathcal{L}(\Box, \mathbf{I})$ is equivalent, over X, to a closed formula in $\mathcal{L}(\Box)$.*

Proof. This is by induction on the closed formula C. The only non-trivial case is $C \equiv \mathbf{I}B$, where B is a closed formula in $\mathcal{L}(\mathbf{I})$. Now, by the induction hypothesis, B is a closed formula in $\mathcal{L}(\Box)$. Furthermore, $il \vdash \mathbf{I}B \leftrightarrow \mathbf{I}(B \vee \Diamond B)$. So, $il \vdash \mathbf{I}B \leftrightarrow \mathbf{I}\Diamond^k\top$, for some $k \in \omega \cup \{\,\omega\,\}$. If $k = 0$, then $\mathbf{I}\Diamond^k\top \equiv \mathbf{I}\top$, and $X \vdash \mathbf{I}B \leftrightarrow \top$. If $k = \omega$, then $\mathbf{I}\Diamond^k\top \equiv \mathbf{I}\bot$, and

$$\begin{aligned} il \vdash \mathbf{I}\bot &\rightarrow (\neg\Diamond\top \vee \Diamond\bot), \text{ by } I3 \\ &\rightarrow (\Box\bot \vee \Diamond\bot) \\ &\rightarrow \Box\bot \\ &\rightarrow \mathbf{I}\bot, \text{ by 1.4.} \end{aligned}$$

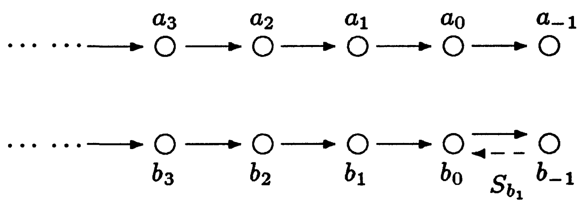
So $X \vdash C \leftrightarrow \Box\bot$. If $0 < k < \omega$, then

$$\begin{aligned} X \vdash \mathbf{I}\Diamond^k\top &\rightarrow \mathbf{I}\Diamond\top, \text{ by axiom } \Box A \rightarrow \Box\Box A \\ &\rightarrow \Box\bot, \text{ by axiom } f \\ &\rightarrow \Box\Diamond^k\top \\ &\rightarrow \mathbf{I}\Diamond^k\top, \text{ by 1.4} \end{aligned}$$

So $X \vdash C \leftrightarrow \Box\bot$. QED.

By the Normal Form Theorem for closed formulas in $\mathcal{L}(\Box)$ it follows from 3.4 that in extensions of $il + f$ every closed formula in $\mathcal{L}(\Box, \mathbf{I})$ is equivalent to a Boolean combination of formulas of the form $\Box^n\bot$, for some $n \in \omega \cup \{\,\omega\,\}$.

Below $il + f$ the situation is more complicated. Note for example that there are infinitely many pairwise non-equivalent closed $\mathcal{L}(\Box, \mathbf{I})$-formulas, none of which is equivalent to a (closed) formula in $\mathcal{L}(\Box)$. To see this, let $A_1 := \mathbf{I}\Diamond\top$, $A_{n+1} := \Diamond(A_n \wedge \Diamond^{n+1}\top)$, and consider the following Veltman-frame $\mathcal{F}$:



Let $\Vdash$ be any forcing relation on $\mathcal{F}$ with, for all $i \in \omega \cup \{-1\}$, $a_i \Vdash p$ iff $b_i \Vdash p$; then for all $B \in \mathcal{L}(\Box)$, $a_i \Vdash B$ iff $b_i \Vdash B$. On the other hand, we have for all $i \in \omega \setminus \{0\}$, $a_i \nVdash A_i$ and $b_i \Vdash A_i$. This shows that none of the A_is is equivalent to an $\mathcal{L}(\Box)$-formula. To see that $il \nvdash A_i \leftrightarrow A_j$, if $i \neq j$, note that for all i, and all $j > i$, $b_i \Vdash A_i \wedge \neg A_j$.

It is still open whether there exist reasonable normal forms for closed formulas in subsystems of $il + f$.

We now examine the modalities in $\mathcal{L}(\mathbf{I})$ and $\mathcal{L}(\Box, \mathbf{I})$. (Recall that a modality is nothing but a sequence consisting of modal operators and/or dual versions of these operators.) We say that two modalities α and β are *equivalent over ils* if for all $A \in \mathcal{L}(\Box, \mathbf{I})$, $ils \vdash \alpha A \leftrightarrow \beta A$. A modality α is called a *constant* modality (over ils) if there is a closed formula C such that for all A, $ils \vdash \alpha A \leftrightarrow C$ (i.e., if for all A, B, $ils \vdash \alpha A \leftrightarrow \alpha B$). We use $\bar{\mathbf{I}}$ as an abbreviation for $\neg\mathbf{I}\neg$.

We start with the modalities over extensions of il. Unlike modalities in more traditional modal languages almost all modalities in $\mathcal{L}(\mathbf{I})$ are constant. For example:

Proposition 3.5 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then*

1. $il \vdash \mathbf{II}A \leftrightarrow \top$;
2. $il \vdash \overline{\mathbf{I}}\overline{\mathbf{I}}A \leftrightarrow \bot$;
3. $il \vdash \mathbf{I}\Box A \leftrightarrow \top$;
4. $il \vdash \bar{\mathbf{I}}\Diamond A \leftrightarrow \bot$.

Proposition 3.6 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then $il \vdash \mathbf{I}\bar{\mathbf{I}}\mathbf{I}A \leftrightarrow \mathbf{I}\bar{\mathbf{I}}\top\ (\leftrightarrow \mathbf{I}\Diamond\top)$.*

Proof. One direction is almost immediate:

$$\begin{aligned} il \vdash \mathbf{I}\bar{\mathbf{I}}\mathbf{I}A \quad &\rightarrow \quad \mathbf{I}\bar{\mathbf{I}}\mathbf{I}\top \\ &\rightarrow \quad \mathbf{I}\bar{\mathbf{I}}\top, \text{ since } il \vdash \Box(\mathbf{I}\top \leftrightarrow \top). \end{aligned}$$

To prove the other one, we show that $il \vdash \bar{\mathbf{I}}\bar{\mathbf{I}}\mathbf{I}A \rightarrow \bar{\mathbf{I}}\mathbf{I}\top$:

$$\begin{aligned} il \vdash \bar{\mathbf{I}}\bar{\mathbf{I}}\mathbf{I}A \wedge \neg\bar{\mathbf{I}}\Box\bot \quad &\rightarrow \quad \mathbf{I}\Diamond\top \wedge \bar{\mathbf{I}}\bar{\mathbf{I}}\mathbf{I}\top \\ &\rightarrow \quad \mathbf{I}\Diamond\top \wedge \bar{\mathbf{I}}\mathbf{I}\Diamond\top, \text{ since } il \vdash \Box(\bar{\mathbf{I}}\top \leftrightarrow \Diamond\top) \\ &\rightarrow \quad \mathbf{I}\Diamond\top \wedge \bar{\mathbf{I}}(\Diamond\top \rightarrow \Diamond\Diamond\top), \text{ by axiom } I4. \end{aligned}$$

Now $il \vdash \mathbf{I}\Diamond\top \wedge \bar{\mathbf{I}}(\Diamond\top \rightarrow \Diamond\Diamond\top) \rightarrow \bot$, by 1.4, and $il \vdash \bar{\mathbf{I}}\Box\bot \leftrightarrow \bar{\mathbf{I}}\mathbf{I}\bot$. Therefore $il \vdash \bar{\mathbf{I}}\bar{\mathbf{I}}\mathbf{I}A \rightarrow \bar{\mathbf{I}}\mathbf{I}\bot$. QED.

As a corollary we find the following result:

Proposition 3.7 *Let X be a logic that extends il. Then*

1. *every modality in $\mathcal{L}(\mathbf{I})$ is equivalent (over X) to one of* $\langle\,\rangle$, $\mathbf{I}$, $\overline{\mathbf{I}}$, $\mathbf{II}$, $\mathbf{I}\overline{\mathbf{I}}$, $\overline{\mathbf{I}}\mathbf{I}$, $\overline{\mathbf{II}}$, $\mathbf{I}\overline{\mathbf{I}}\mathbf{I}$, $\mathbf{I}\overline{\mathbf{II}}$, $\overline{\mathbf{I}}\mathbf{II}$ *or* $\overline{\mathbf{I}}\mathbf{I}\overline{\mathbf{I}}$;
2. *if X is il then the only non-constant modalities in $\mathcal{L}(\mathbf{I})$ are* $\langle\,\rangle$, $\mathbf{I}$, $\overline{\mathbf{I}}$, $\mathbf{I}\overline{\mathbf{I}}$ *and* $\overline{\mathbf{I}}\mathbf{I}$.

Proof. Note first that if α, β are one of the modalities mentioned in item 1, and if $\alpha \not\equiv \beta$, then α and β are not equivalent over *il*. Let α be a modality in $\mathcal{L}(\mathbf{I})$. Then either $\alpha \in \{\langle\,\rangle, \mathbf{I}, \overline{\mathbf{I}}, \mathbf{I}\overline{\mathbf{I}}, \overline{\mathbf{I}}\mathbf{I}\}$, and we are done, or for some α' we have $\alpha \in \{\mathbf{II}\alpha', \mathbf{I}\overline{\mathbf{I}}\mathbf{I}\alpha', \overline{\mathbf{II}}\mathbf{I}\alpha', \overline{\mathbf{II}}\alpha', \overline{\mathbf{I}}\mathbf{II}\alpha', \overline{\mathbf{I}}\mathbf{I}\overline{\mathbf{I}}\alpha'\}$. In the latter case an application of 3.5 or 3.6 yields item 1.

To prove item 2, note first that $\langle\,\rangle$, $\mathbf{I}$, $\overline{\mathbf{I}}$, $\mathbf{I}\overline{\mathbf{I}}$ and $\overline{\mathbf{I}}\mathbf{I}$ are indeed non-constant modalities; that they are the only such modalities in $\mathcal{L}(\mathbf{I})$ is immediate from 3.5, 3.6 and item 1. QED.

Proposition 3.8 *Let X be a logic that extends il. Then every modality in $\mathcal{L}(\Box, \mathbf{I})$ is equivalent (over X) to a modality of the form $\alpha_1\beta_1 \ldots \alpha_n\beta_n$, where the α_is are modalities in $\mathcal{L}(\Box)$ and for $1 \leq i < n$, $\beta_i \in \{\langle\,\rangle, \mathbf{I}, \overline{\mathbf{I}}, \mathbf{I}\overline{\mathbf{I}}, \overline{\mathbf{I}}\mathbf{I}\}$, while $\beta_n \in \{\langle\,\rangle, \mathbf{I}, \overline{\mathbf{I}}, \mathbf{II}, \mathbf{I}\overline{\mathbf{I}}, \overline{\mathbf{I}}\mathbf{I}, \overline{\mathbf{II}}, \mathbf{I}\overline{\mathbf{I}}\mathbf{I}, \mathbf{I}\overline{\mathbf{II}}, \overline{\mathbf{I}}\mathbf{II}, \overline{\mathbf{I}}\mathbf{I}\overline{\mathbf{I}}\}$.*

We continue with a somewhat simpler case: the modalities over extensions of *ilm*. Here there are even fewer non-constant modalities in $\mathcal{L}(\mathbf{I})$. For a start, we have the following stronger version of 3.6.

Proposition 3.9 *Let A, $B \in \mathcal{L}(\Box, \mathbf{I})$. Then $ilm \vdash \mathbf{I}\overline{\mathbf{I}}A \leftrightarrow \Box\bot$.*

Proof. Since $ilm \vdash \Box\bot \rightarrow \Box\overline{\mathbf{I}}A$, we have $ilm \vdash \Box\bot \rightarrow \mathbf{I}\overline{\mathbf{I}}A$, by 1.4. To prove the converse, note that $ilm \vdash \Box(\overline{\mathbf{I}}A \wedge \Box\bot \rightarrow \bot)$. So since $ilm \vdash \mathbf{I}\overline{\mathbf{I}}A \rightarrow \mathbf{I}(\overline{\mathbf{I}}A \wedge \Box\bot)$, by axiom m, we have $ilm \vdash \mathbf{I}\overline{\mathbf{I}}A \rightarrow \mathbf{I}\bot$, by axiom $I2$. Thus $ilm \vdash \mathbf{I}\overline{\mathbf{I}}A \rightarrow \Box\bot$. QED.

Proposition 3.10 *Let X be a logic that extends ilm. Then every modality in $\mathcal{L}(\mathbf{I})$ is equivalent (over X) to one of* $\langle\,\rangle$, $\mathbf{I}$, $\overline{\mathbf{I}}$, $\mathbf{II}$, $\mathbf{I}\overline{\mathbf{I}}$, $\overline{\mathbf{I}}\mathbf{I}$ *or* $\overline{\mathbf{II}}$. *Moreover, if X is ilm or ilp then the only non-constant modalities in $\mathcal{L}(\mathbf{I})$ are* $\langle\,\rangle$, $\mathbf{I}$ *and* $\overline{\mathbf{I}}$.

Proof. Immediate from 3.5 and 3.9. QED.

Proposition 3.11 *Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then*

1. $ilm \vdash \mathbf{I}\Diamond A \leftrightarrow \Box\bot$;
2. $ilm \vdash \overline{\mathbf{I}}\Box A \leftrightarrow \Diamond\top$.

Proposition 3.12 *Let X be a logic that extends ilm. Then*

1. *every modality in $\mathcal{L}(\Box, \mathbf{I})$ is equivalent (over X) to a modality of the form $\alpha\beta$, where α is a (possibly empty) modality in $\mathcal{L}(\Box)$, and $\beta \in \{\langle\,\rangle, \mathbf{I}, \overline{\mathbf{I}}, \mathbf{I}\overline{\mathbf{I}}, \overline{\mathbf{I}}\mathbf{I}, \mathbf{II}, \overline{\mathbf{II}}, \mathbf{I}\Diamond, \overline{\mathbf{I}}\Box\}$;*
2. *if X is ilm or ilp, then the only non-constant modalities in $\mathcal{L}(\Box, \rhd)$ are $\Diamond^k$, $\Box^k$, $\Box^k\mathbf{I}$ and $\Diamond^k\overline{\mathbf{I}}$.*

Proof. We only prove item 2. Let γ be a non-constant modality in $\mathcal{L}(\Box, \rhd)$; by item 1 we may assume that $\gamma \equiv \alpha\beta$, with $\alpha\beta$ as described in item 1. Since γ is assumed to be non-constant, $\beta \in \{\langle\,\rangle, \mathbf{I}, \overline{\mathbf{I}}\}$. Moreover since $\Box\Diamond$ and $\Diamond\Box$ are constant we may assume that $\alpha \equiv \Diamond^k$, or $\alpha \equiv \Box^k$, for some k.

If $\beta \equiv \langle\,\rangle$, then $\gamma \equiv \Diamond^k$ or $\gamma \equiv \Box^k$; in both cases γ is non-constant for all k.

If $\beta \equiv \mathbf{I}$, then $\gamma \equiv \Diamond^k\mathbf{I}$ or $\gamma \equiv \Box^k\mathbf{I}$. Since $ilm \vdash \Diamond\mathbf{I}A \leftrightarrow \Diamond\top$, we have that $\Diamond^k\mathbf{I}$ is constant for all $k \geq 1$; on the other hand, for any k, $\Box^k\mathbf{I}$ is non-constant, as the reader may verify.

Similarly, if $\beta \equiv \bar{\mathbf{I}}$, then γ is non-constant iff $\gamma \equiv \Diamond^k\bar{\mathbf{I}}$. QED.

For the remainder of this section let T be a Σ^0_1-sound essentially reflexive theory. (Modulo some obvious changes most of the remarks in the sequel hold equally well for Σ^0_1-sound finitely axiomatized sequential theories that extend $I\Delta_0 + \mathrm{SupExp}$.) Let $\Box_T$ be a formalization (in the language of T) of provability in T; $\Diamond_T\varphi$ is short for $\neg\Box_T\neg\varphi$; $\mathbf{I}_T$ is a formalization (in the language of T) of the unary interpretability predicate over T.

Assume that φ is a sentence in the language of T that is not of the form $(\neg)\mathbf{I}_T\psi$ or $(\neg)\Box_T\psi$. We want to know what the theory T can say about sentences of the form $\beta\varphi$, where β is (the arithmetical version of) a non-empty modality of the form $(\neg)\mathbf{I}\beta'$. By 3.12 we only have to consider 8 cases.

Note first that no formula of the form $\neg\mathbf{I}_T\varphi$ can be provable in T. For, we have $ilm \vdash \Box\neg\mathbf{I}A \to \Box\bot$, for all $A \in \mathcal{L}(\Box,\mathbf{I})$. So $T \vdash \Box_T\neg\mathbf{I}_T\varphi \to \Box_T(0=1)$, for all sentences φ in the language of T. Therefore, if $T \vdash \neg\mathbf{I}_T\varphi$ then $T \vdash \Box_T(0=1)$. Since T is assumed to be Σ^0_1-sound, this implies that for no φ, $T \vdash \neg\mathbf{I}_T\varphi$.

Similarly, since $ilm \vdash \mathbf{I}\Diamond A \leftrightarrow \Box\bot$ and $ilm \vdash \mathbf{I}\bar{\mathbf{I}}A \leftrightarrow \Box\bot$, we can not have $T \vdash \mathbf{I}_T\Diamond_T\varphi$ or $T \vdash \mathbf{I}_T\bar{\mathbf{I}}_T\varphi$, for any sentence φ. Moreover, we do have for all sentences φ, $T \vdash \mathbf{I}_T\mathbf{I}_T\varphi$, because $ilm \vdash \mathbf{I}\mathbf{I}A$. The only remaining case, then, is $\beta \equiv \mathbf{I}$. Here we have the following possibilities:

1. $T \vdash \varphi$, and then $T \vdash \mathbf{I}_T\varphi$, $T \nvdash \mathbf{I}_T\neg\varphi$;
2. $T \vdash \neg\varphi$, and then $T \nvdash \mathbf{I}_T\varphi$, $T \vdash \mathbf{I}_T\neg\varphi$;
3. $T \nvdash \varphi$, $T \nvdash \neg\varphi$ and $T \vdash \mathbf{I}_T\varphi$, $T \vdash \mathbf{I}_T\neg\varphi$;
4. $T \nvdash \varphi$, $T \nvdash \neg\varphi$ and $T \vdash \mathbf{I}_T\varphi$, $T \nvdash \mathbf{I}_T\neg\varphi$;
5. $T \nvdash \varphi$, $T \nvdash \neg\varphi$ and $T \nvdash \mathbf{I}_T\varphi$, $T \vdash \mathbf{I}_T\neg\varphi$;
6. $T \nvdash \varphi$, $T \nvdash \neg\varphi$ and $T \nvdash \mathbf{I}_T\varphi$, $T \nvdash \mathbf{I}_T\neg\varphi$.

By our previous remarks no strengthening of this classification is possible by replacing '$T \nvdash$ ' by '$T \vdash \neg$' somewhere.

We leave it to the reader to supply examples of items 1 and 2; the sentence $\Box_T(0=1)$ is a sentence that satisfies item 4, and its negation satisfies item 5; below we will provide examples of sentences that satisfy items 3 and 6, respectively. Recall that an *Orey sentence* for T is a sentence ψ such that both ψ and $\neg\psi$ are interpretable in T. So a sentence satisfying item 3 is an example of a sentence that is provably in T an Orey sentence for T. Our example below of a sentence satisfying item 6 is also an example of a sentence that is—unprovably in T—an Orey sentence for T.

Example 3.13 *There is a sentence φ that satisfies item 3.*

Proof. Put $A \equiv \neg\Box p \wedge \neg\Box\neg p \wedge \Box\mathbf{I}p \wedge \Box\mathbf{I}\neg p$. We prove that $ilm^\omega \nvdash \neg A$; then, by 2.18, there is an interpretation $(\cdot)^T$ of $\mathcal{L}(\Box,\mathbf{I})$ in the language of T such that $(\neg A)^T$ is false in the standard model. Hence $(A)^T$ is true. Put $\varphi = (p)^T$ and we are done.

Now, to prove that $ilm^\omega \nvdash \neg A$ we show that

$$ilm \nvdash \Big(\bigwedge_{\Box B \in \mathrm{Sub}(\neg A)} (\Box B \to B) \wedge \bigwedge_{\mathbf{I}D \in \mathrm{Sub}(\neg A)} \Diamond \top \Big) \to \neg A. \qquad (3)$$

Define $\mathcal{M}$ as in Figure 4:

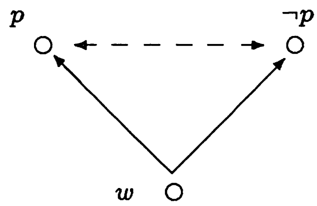


Figure 4.

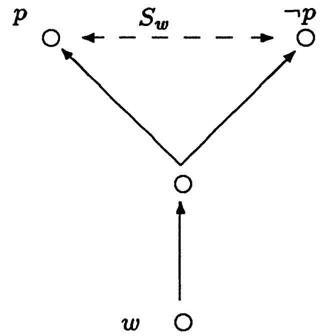


Figure 5.

We leave it to the reader to check that $w \Vdash \bigwedge_{\mathbf{I}D \in \mathrm{Sub}(\neg A)} \Diamond \top$ and that $w \Vdash \bigwedge_{\Box B \in \mathrm{Sub}(\neg A)} (\Box B \to B)$; from this and $w \Vdash A$ we obtain (3). QED.

Example 3.14 *There is a sentence φ that satisfies item 6, and such that φ is, unprovably in T, an Orey sentence for T.*

Proof. Put $A \equiv \neg\Box p \wedge \neg\Box\neg p \wedge \neg\Box\mathbf{I}p \wedge \neg\Box\mathbf{I}\neg p \wedge \mathbf{I}p \wedge \mathbf{I}\neg p$. We only have to show that $ilm^\omega \nvdash \neg A$, then we find an interpretation $(\cdot)^T$ of $\mathcal{L}(\Box, \mathbf{I})$ in the language of T such that $(A)^T$ is true. Put $\varphi = (p)^T$ and we are done.

We leave it to the reader to check that the model depicted in Figure 5 shows that $ilm^\omega \nvdash \neg A$. QED.

Note that the model used in 3.14 is not an ILP-model. Therefore, the sentence φ given there works only for essentially reflexive theories T. We leave it to the reader to find a φ that satisfies item 6 if T is a Σ^0_1-sound finitely axiomatized theory that extends $\mathrm{I}\Delta_0 + \mathrm{SupExp}$. He or she won't be able to find a sentence φ that satisfies 3.14 for such T. For, let T be such a theory, and assume that $T \nvdash \mathbf{I}_T\varphi$ while $T + \varphi$ is interpretable in T. Then $\omega \models \mathbf{I}_T\varphi$. Hence, $\omega \models \Box_T\mathbf{I}_T\varphi$ (since $\omega \models \mathbf{I}_T\varphi \to \Box_T\mathbf{I}_T\varphi$), and so $T \vdash \mathbf{I}_T\varphi$—a contradiction.

An inspection of the arithmetical completeness proof of ILM shows that the sentences φ found in 3.13 and 3.14 may be taken to be Σ^0_2-sentences.

3.2 Interpolation and Fixed Point Theorems

Our proof of the interpolation theorem for *il*, *ilm* and *ilp* extends Smoryński's proof of the interpolation theorem for L. (Cf. [5].)

Definition 3.15 Let $A \in \mathcal{L}(\Box, \mathbf{I})$. Then $\mathcal{L}_A$ is the sublanguage of $\mathcal{L}(\Box, \mathbf{I})$ consisting of all formulas having only proposition letters occurring in A. A set

$X \subseteq \mathcal{L}_A$ is maximal *ils*-consistent *in* $\mathcal{L}_A$ if for all $C \in \mathcal{L}_A$, either $C \in \mathcal{L}_A$ or $\neg C \in \mathcal{L}_A$.

A pair $\langle X, Y\rangle$ with $X \subseteq \mathcal{L}_A$, $Y \subseteq \mathcal{L}_B$ is called *separable* if for some $C \in \mathcal{L}_A \cap \mathcal{L}_B$, $C \in X$ and $\neg C \in Y$. If $\langle X, Y\rangle$ is not separable it is *inseparable*.

A pair $\langle X, Y\rangle$ with $X \subseteq \mathcal{L}_A$, $Y \subseteq \mathcal{L}_B$ is called a *complete* pair if

1. $\langle X, Y\rangle$ is separable.
2. X is maximal *ils*-consistent in $\mathcal{L}_A$.
3. Y is maximal *ils*-consistent in $\mathcal{L}_B$.

Our proof of the interpolation theorem for *il* (*ilm*, *ilp*) is in fact nothing but another modal completeness proof for *il* (*ilm*, *ilp*)—using complete pairs instead of plain maximal *il* (*ilm*, *ilp*)-consistent sets. The construction of a counter model is entirely analogous to the constructions in 2.6, 2.11 and 2.21. The main difference is the result that supplies us with the input for our construction. That is: 2.3, 2.4 and 2.5 have to be restated and reproved for complete pairs.

Definition 3.16 Let $\langle X, Y\rangle$, $\langle X', Y'\rangle$ be complete pairs.

1. $\langle X, Y\rangle \prec \langle X', Y'\rangle$ ($\langle X', Y'\rangle$ is a *successor* of $\langle X, Y\rangle$) if
   (a) $A \in X' \cup Y'$ for all $\Box A \in X \cup Y$
   (b) $\Box A \in X' \cup Y'$ for some $\Box A \notin X \cup Y$
2. $\langle X', Y'\rangle$ is called a *C-critical successor* of $\langle X, Y\rangle$ if
   (a) $\langle X, Y\rangle \prec \langle X', Y'\rangle$
   (b) $\mathbf{I}C \notin X \cup Y$
   (c) $\neg C, \Box\neg C \in X' \cup Y'$

Proposition 3.17 *Let $X_0 \subseteq \mathcal{L}_A$, $Y_0 \subseteq \mathcal{L}_B$ be such that $\langle X_0, Y_0\rangle$ is an inseparable pair. Then there exists a complete pair $\langle X, Y\rangle$ with $X_0 \subseteq X \subseteq \mathcal{L}_A$ and $Y_0 \subseteq Y \subseteq \mathcal{L}_B$.*

Proof. See [5], Lemma 1.1. QED.

Proposition 3.18 *Let $\langle X, Y\rangle$ be a complete pair such that $\Diamond C \in X \cup Y$. Then there exists a complete pair $\langle X', Y'\rangle \succ \langle X, Y\rangle$ with $C, \Box\neg C \in X' \cup Y'$.*

Proof. See [5], Lemma 1.2. QED.

Proposition 3.19 *Let $\langle X, Y\rangle$ be a complete pair such that $\mathbf{I}C \notin X \cup Y$. Then there exists a C-critical complete pair $\langle X', Y'\rangle \succ \langle X, Y\rangle$ with $\Box\bot \in X' \cup Y'$.*

Proof. Assume that no such $\langle X', Y'\rangle$ exists. We distinguish 3 cases.
Case 1. $\mathbf{I}C \in \mathcal{L}_A \backslash \mathcal{L}_B$. Then by 3.17 and compactness there are $\Box F_1, \dots, \Box F_m \in X$, $\Box G_1, \dots, \Box G_n \in Y$ and $D \in \mathcal{L}_A \cap \mathcal{L}_B$ such that

$$F_1, \dots, F_m, \neg C, \Box\neg C, \Box\bot \vdash D \quad (4)$$
$$G_1, \dots, G_n, \Box\bot \vdash \neg D. \quad (5)$$

By (4) we have $\Box F_1, \dots, \Box F_m \vdash \Box(\Box\bot \to (\neg D \to C \vee \Diamond C))$. Now

$$il \vdash \neg \mathbf{I}C \wedge \Box(\Box\bot \to (\neg D \to C \vee \Diamond C)) \to \neg\Box(\Box\bot \to \neg D).$$

So $X \vdash \neg\Box(\Box\bot \to \neg D)$. On the other hand, (5) yields $Y \vdash \Box(\Box\bot \to \neg D)$. So X and Y are separable—a contradiction.

Case 2. $\mathbf{I}C \in \mathcal{L}_B \setminus \mathcal{L}_A$. Similar to Case 1.
Case 3. $\mathbf{I}C \in \mathcal{L}_A \cap \mathcal{L}_B$. By 3.17 and compactness there exist $\Box F_1, \ldots, \Box F_m \in X$, $\Box G_1, \ldots, \Box G_n \in Y$ and $D \in \mathcal{L}_A \cap \mathcal{L}_B$ such that

$$F_1, \ldots, F_m, \neg C, \Box\neg C, \Box\bot \vdash D \quad (6)$$
$$G_1, \ldots, G_n, \neg C, \Box\neg C, \Box\bot \vdash \neg D. \quad (7)$$

Now $il \vdash \neg\mathbf{I}C \wedge \Box(\Box\bot \to (\neg D \to C \vee \Diamond C)) \to \neg\Box(\Box\bot \to (D \to C \vee \Diamond C))$. So (6) yields

$$\Box F_1, \ldots, \Box F_m \vdash \Box(\Box\bot \to (\neg D \to C \vee \Diamond C))$$
$$\Box F_1, \ldots, \Box F_m, \neg\mathbf{I}C \vdash \neg\Box(\Box\bot \to (D \to C \vee \Diamond C))$$
$$X \vdash \neg\Box(\Box\bot \to (D \to C \vee \Diamond C)).$$

On the other hand (7) gives $Y \vdash \Box(\Box\bot \to (D \to C \vee \Diamond C))$. Hence X and Y are separable—a contradiction. QED.

Proposition 3.20 *Let $\langle X, Y\rangle$ be a complete pair with $\neg\mathbf{I}C \in X \cup Y$ and $\mathbf{I}E \in X \cup Y$. Then there exists a C-critical complete pair $\langle X', Y'\rangle \succ \langle X, Y\rangle$ with $E, \Box\neg E \in X' \cup Y'$.*

Proof. Assume that no such $\langle X', Y'\rangle$ exists. We distinguish 9 cases.
Case 1. $\mathbf{I}E \in \mathcal{L}_A \setminus \mathcal{L}_B$, $\mathbf{I}C \in \mathcal{L}_A \setminus \mathcal{L}_B$. By 3.17 and compactness there exist $\Box F_1, \ldots, \Box F_m \in X$, $\Box G_1, \ldots, \Box G_n \in Y$ and $D \in \mathcal{L}_A \cap \mathcal{L}_B$ such that

$$F_1, \ldots, F_m, \neg C, \Box\neg C, E, \Box\neg E \vdash D \quad (8)$$
$$G_1, \ldots, G_n \vdash \neg D. \quad (9)$$

Now (8) yields

$$\Box F_1, \ldots, \Box F_m, \Box\neg D \vdash \Box(E \wedge \Box E \to C \vee \Diamond C)$$
$$\Box F_1, \ldots, \Box F_m, \Box\neg D \vdash \mathbf{I}(E \wedge \Box\neg E) \to \mathbf{I}(C \vee \Diamond C), \text{ by axiom } I2$$
$$\Box F_1, \ldots, \Box F_m, \Box\neg D \vdash \mathbf{I}E \to \mathbf{I}C, \text{ by 1.4.(3) and axiom } I3$$
$$\Box F_1, \ldots, \Box F_m \vdash \mathbf{I}E \wedge \neg\mathbf{I}C \to \neg\Box\neg D$$
$$X \vdash \neg\Box\neg D.$$

On the other hand (9) yields $Y \vdash \Box\neg D$. So X and Y are separable—a contradiction.
Case 2. $\mathbf{I}E \in \mathcal{L}_A \setminus \mathcal{L}_B$, $\mathbf{I}C \in \mathcal{L}_B \setminus \mathcal{L}_A$. Then by 3.17 and compactness there exist $\Box F_1, \ldots, \Box F_m \in X$, $\Box G_1, \ldots, \Box G_n \in Y$ and $D \in \mathcal{L}_A \cap \mathcal{L}_B$ such that

$$F_1, \ldots, F_m, E, \Box\neg E \vdash D \quad (10)$$
$$G_1, \ldots, G_n, \neg C, \Box\neg C \vdash \neg D \quad (11)$$

As before (10) yields

$$\Box F_1, \ldots, \Box F_m \vdash \Box(E \wedge \Box\neg E \to D)$$
$$\Box F_1, \ldots, \Box F_m \vdash \mathbf{I}(E \wedge \Box\neg E) \to \mathbf{I}D$$
$$\Box F_1, \ldots, \Box F_m, \mathbf{I}E \vdash \mathbf{I}D$$
$$X \vdash \mathbf{I}D.$$

But (11) yields

$$\begin{aligned}\Box G_1,\ldots,\Box G_n &\vdash \Box(D \to C \vee \Diamond C)\\ \Box G_1,\ldots,\Box G_n, \mathbf{I}D &\vdash \mathbf{I}C, \text{by axioms } I2 \text{ and } I3\\ \Box G_1,\ldots,\Box G_n, \neg\mathbf{I}C &\vdash \neg\mathbf{I}D\\ Y &\vdash \neg\mathbf{I}D.\end{aligned}$$

And X and Y are separable after all—a contradiction.

Case 3. $\mathbf{I}E \in \mathcal{L}_A \setminus \mathcal{L}_B$, $\mathbf{I}C \in \mathcal{L}_A \cap \mathcal{L}_B$. Then by 3.17 and compactness there exist $\Box F_1,\ldots,\Box F_m \in X$, $\Box G_1,\ldots,\Box G_n \in Y$ and $D \in \mathcal{L}_A \cap \mathcal{L}_B$ such that

$$F_1,\ldots,F_m, \neg C, \Box\neg C, E, \Box\neg E \vdash D \tag{12}$$

$$G_1,\ldots,G_n, \neg C, \Box\neg C \vdash \neg D. \tag{13}$$

By (12) we find $\Box F_1,\ldots,\Box F_m \vdash \Box(E \wedge \Box\neg E \to (\neg D \to C \vee \Diamond C))$. Now

$$il \vdash \neg\mathbf{I}C \wedge \mathbf{I}E \wedge \Box(E \wedge \Box\neg E \to (\neg D \to C \vee \Diamond C)) \to \neg\Box(D \to C \vee \Diamond C),$$

so $X \vdash \neg\Box(D \to C \vee \Diamond C)$. On the other hand, (13) yields $Y \vdash \Box(D \to C \vee \Diamond C)$. Again, this implies that X and Y are separable—a contradiction.

Case 4. $\mathbf{I}E \in \mathcal{L}_B \setminus \mathcal{L}_A$, $\mathbf{I}C \in \mathcal{L}_B \setminus \mathcal{L}_A$. Similar to Case 1.

Case 5. $\mathbf{I}E \in \mathcal{L}_B \setminus \mathcal{L}_A$, $\mathbf{I}C \in \mathcal{L}_A \setminus \mathcal{L}_B$. Similar to Case 2.

Case 6. $\mathbf{I}E \in \mathcal{L}_B \setminus \mathcal{L}_A$, $\mathbf{I}C \in \mathcal{L}_A \cap \mathcal{L}_B$. Similar to Case 3.

Case 7. $\mathbf{I}E \in \mathcal{L}_A \cap \mathcal{L}_B$, $\mathbf{I}C \in \mathcal{L}_A \setminus \mathcal{L}_B$. Then by 3.17 and compactness there are $\Box F_1,\ldots,\Box F_m \in X$, $\Box G_1,\ldots,\Box G_n \in Y$ and $D \in \mathcal{L}_A \cap \mathcal{L}_B$ such that

$$F_1,\ldots,F_m, \neg C, \Box\neg C, E, \Box\neg E \vdash D \tag{14}$$

$$G_1,\ldots,G_n, E, \Box\neg E \vdash \neg D. \tag{15}$$

By (14) we have $\Box F_1,\ldots,\Box F_m \vdash \Box(E \wedge \Box\neg E \to (\neg D \to C \vee \Diamond C))$. Now

$$il \vdash \neg\mathbf{I}C \wedge \mathbf{I}E \wedge \Box(E \wedge \Box\neg E \to (\neg D \to C \vee \Diamond C)) \to \neg\Box(E \wedge \Box\neg E \to \neg D),$$

so $X \vdash \neg\Box(E \wedge \Box\neg E \to \neg D)$. On the other hand, (15) yields $Y \vdash \Box(E \wedge \Box\neg E \to \neg D)$. So X and Y are separable—a contradiction.

Case 8. $\mathbf{I}E \in \mathcal{L}_A \cap \mathcal{L}_B$, $\mathbf{I}C \in \mathcal{L}_B \setminus \mathcal{L}_A$. Similar to Case 7.

Case 9. $\mathbf{I}E \in \mathcal{L}_A \cap \mathcal{L}_B$, $\mathbf{I}C \in \mathcal{L}_A \cap \mathcal{L}_B$. Then by 3.17 and compactness there exist $\Box F_1,\ldots,\Box F_m \in X$, $\Box G_1,\ldots,\Box G_n \in Y$ and $D \in \mathcal{L}_A \cap \mathcal{L}_B$ such that

$$F_1,\ldots,F_m, \neg C, \Box\neg C, E, \Box\neg E \vdash D \tag{16}$$

$$G_1,\ldots,G_n, \neg C, \Box\neg C, E, \Box\neg E \vdash \neg D. \tag{17}$$

Now $il \vdash \neg\mathbf{I}C \wedge \mathbf{I}E \wedge \Box(E \wedge \Box\neg E \to (\neg D \to C \vee \Diamond C)) \to \neg\Box(E \wedge \Box\neg E \to (D \to C \vee \Diamond C))$, so (16) yields

$$\begin{aligned}\Box F_1,\ldots,\Box F_m, \mathbf{I}E, \neg\mathbf{I}C &\vdash \neg\Box(E \wedge \Box\neg E \to (D \to C \vee \Diamond C))\\ X &\vdash \neg\Box(E \wedge \Box\neg E \to (D \to C \vee \Diamond C)).\end{aligned}$$

But (17) yields $Y \vdash \Box(E \wedge \Box\neg E \to (D \to C \vee \Diamond C))$. And again, X and Y are separable—a contradiction. QED.

Theorem 3.21 (Interpolation Theorem) *Let ils be one of il, ilm or ilp. If ils* $\vdash A \to B$*, then there is a formula* C *having only proposition letters occurring in both* A *and* B *such that ils* $\vdash A \to C$ *and ils* $\vdash C \to B$.

Proof. The proof is by contraposition. Fix A and B and assume that no interpolant exists. We will show that $ils \nvdash A \to B$ by constructing a counter model to the implication.

Note that the assumption that no interpolant between A and B exists, means: $\{A\}$ and $\{\neg B\}$ are separable. So by 3.17 there exists a complete pair $\langle X, Y\rangle$ with $\{A\} \subseteq X \subseteq \mathcal{L}_A$ and $\{\neg B\} \subseteq Y \subseteq \mathcal{L}_B$.

Put $\Gamma := \langle X, Y\rangle$ and construct W_Γ as in 2.6 (or 2.11 if $ils = ilm$, and 2.21 if $ils = ilp$)—starting with $\langle \Gamma, \langle\langle\rangle\rangle\rangle$ and adding pairs $\langle \Delta, \tau\rangle$ consisting of complete pairs Δ and sequences τ of pairs (or triples) of formulas. Using 3.18, 3.19 and 3.20 one can then mimic the proof of 2.9 (or 2.14 or 2.23) to find a counter model to the implication $A \to B$. QED.

To state Beth's Theorem and the Fixed Point Theorem for *il*, *ilm* and *ilp*, we first introduce some notation and terminology. We use $A(p)$ for a formula in which p possibly occurs; p is said to occur *modalized* in $A(p)$ if p occurs only in the scope of a $\Box$ or a $\mathbf{I}$. $A(C)$ denotes the result of substituting C for p in $A(p)$.

Theorem 3.22 (Beth's Theorem) *Let $A(r) \in \mathcal{L}(\Box, \mathbf{I})$ contain neither proposition letter p nor q. If $ils \vdash A(p) \wedge A(q) \to (p \leftrightarrow q)$ then, for some $C \in \mathcal{L}_{A(r)} \setminus \{r\}$, $ils \vdash A(p) \to (p \leftrightarrow C)$.*

Proof. The Theorem may be derived from 3.21 in a standard way. Cf. [5]. QED.

Proposition 3.23 *1. $il \vdash \Box(A \leftrightarrow B) \to (\mathbf{I}A \leftrightarrow \mathbf{I}B)$*
2. $il \vdash \Box^+(B \leftrightarrow C) \to (A(B) \leftrightarrow A(C))$.

If p occurs modalized in $A(p)$ and B is a conjunction of formulas of the form $\Box E$ and $\Box^+ E$ then

3. $il \vdash \Box(C \leftrightarrow D) \to (A(C) \leftrightarrow A(D))$
4. $il \vdash B \to (\Box A \to A)$ implies $il \vdash B \to A$
5. $il \vdash \Box^+(p \leftrightarrow A(p)) \wedge \Box^+(q \leftrightarrow A(q)) \to (p \leftrightarrow q)$.

Theorem 3.24 (Explicit Definability of Fixed Points) *Let p occur modalized in $A(p)$. Then there is a formula B with only those proposition letters of A other than p and such that $il \vdash B \leftrightarrow A(B)$.*

Proof. The Theorem may be derived from 3.22 and 3.23 in a standard way. Cf. [5]. QED.

4 Concluding remarks

In [6] the bi-modal provability logic PRL_1 is defined in a modal language $\mathcal{L}(\Box_1, \Box_2)$ with two provability operators. Besides Modus Ponens it has as a rule of inference Necessitation for $\Box_1$; its axioms are the usual L-axioms for $\Box_1$ plus $\Box_2(A \to B) \to (\Box_2 A \to \Box_2 B)$, $\Box_1 A \to \Box_2 A$ and $\Box_2 A \to \Box_1\Box_2 A$. Define a translation $(\cdot)^* : \mathcal{L}(\Box, \mathbf{I}) \longrightarrow \mathcal{L}(\Box_1, \Box_2)$ by

$$\begin{aligned} p^t &:= p \\ (\neg A)^t &:= \neg A^t \\ (A \wedge B)^t &:= A^t \wedge B^t \\ (\Box A)^t &:= \Box_1\Box_2 A^t \\ (\mathbf{I}A)^t &:= \Box_1(\Diamond_2 \top \to \Diamond_2 A^t). \end{aligned}$$

Using Albert Visser's alternative semantics for ILP (cf. [9]) one may then show that for all $A \in \mathcal{L}(\Box, \mathbf{I})$, $ilp \vdash A$ iff $PRL_1 \vdash A^t$.

This much about a connection of (one of) our new logics with a previously known one. Let's look in the opposite direction now, and consider an extension of the language $\mathcal{L}(\Box, \mathbf{I})$. Montagna and Hájek [3] show that ILM is the logic of Π_1^0-conservativity in the following sense: given a Σ_1^0-sound extension T of $I\Sigma_1$ define the interpretation $(A \rhd B)^*$ of a formula $A \rhd B$ in the language of T to be '$T + B^*$ is Π_1^0-conservative over $T + A^*$'; then $ILM \vdash A$ iff for all such $(\cdot)^*$, $T \vdash A^*$. It is well-known that in essentially reflexive theories like PA relative interpretability and Π_1^0-conservativity (in the above sense) are provably extensionally equivalent. However in finitely axiomatized theories like $I\Sigma_1$ the two notions no longer coincide. So it is natural to extend $\mathcal{L}(\Box, \rhd)$ with an operator $\rhd_M$ to be interpreted arithmetically as Π_1^0-conservativity. (It's convenient in this context to write $\rhd_P$ instead of $\rhd$ for the 'old' operator $\rhd$.) As axioms we take the usual L-axioms and rules plus the ILM-axioms for $\rhd_M$, and the ILP-axioms for $\rhd_P$. In addition we have the following 'mixed' axiom: $A \rhd_M B \to A \wedge (C \rhd_P D) \rhd_M B \wedge (C \rhd_P D)$. The resulting system is called ILM/P. The relevant models are tuples $\langle W, R, S^M, S^P, \Vdash\rangle$ where $\langle W, R, S^M, \Vdash\rangle$ is an ILM-model, $\langle W, R, S^P, \Vdash\rangle$ is an ILP-model, while the following extra condition holds: if $xRyS_x^M zRuS_y^P v$ then $uS_z^P v$. It is still open whether ILM/P is modally complete with respect to such ILM/P-models. The unary counterpart ilm/p of ILM/P is defined in a language $\mathcal{L}(\mathbf{I}_M, \mathbf{I}_P)$ with two unary interpretability operators; its axioms and rules are those of L plus the ilm-axioms for $\mathbf{I}_M$ and the ilp-axioms for $\mathbf{I}_P$; ilm/p has no 'mixed' axioms. It has been shown by the present author that ilm/p is modally complete w.r.t. ILM/P-models.

We end with a remark on the method used here to prove modal completeness results for the unary logics. Recall that it employs infinite maximal consistent sets and a 'small' adequate set instead of finite maximal consistent sets that are contained in a 'large' adequate set (as used, for example, in [6] and [2]). Our method has already been used to prove the modal completeness of several of the binary interpretability logics mentioned in this paper—however, it is still open whether ILM may be proved modally complete using this method (cf. [4]).

Acknowledgment

Some of the results listed in this paper were obtained as part of my Master's Thesis written at the Department of Mathematics and Computer Science of the University of Amsterdam under the supervision of Dick de Jongh; I want to thank him for his questions and suggestions.

References

[1] Alessandro Berarducci. The Interpretability Logic of Peano Arithmetic. Manuscript, March 14, 1989.

[2] Dick de Jongh & Frank Veltman. Provability Logics for Interpretability. *ITLI Prepublication Series ML-88-03*, University of Amsterdam. To appear in: *Proceedings of the 1988 Heyting Colloquium*, Plenum Press, Boston, 1990.

[3] F. Montagna & P. Hájek. *ILM* is the Logic of Π^0_1- conservativity. Preprint, Siena, 1989.

[4] Maarten de Rijke. Some Chapters on Interpretability Logic. *ITLI Prepublication Series X-90-02*, University of Amsterdam, 1990.

[5] Craig Smoryński. Beth's Theorem and Self-Referential Statements. In: A. Macintyre et. al. (eds.) *Logic Colloquium '77*. North-Holland, Amsterdam, 1978.

[6] Craig Smoryński. *Self-Reference and Modal Logic*. Springer-Verlag, New York, 1985.

[7] Vítěslav Švejdar. Some Independence Results in Interpretability Logic. Preprint.

[8] Albert Visser. Preliminary Notes on Interpretability Logic. *Logic Group Preprint Series* No. 14, Department of Philosophy, University of Utrecht, 1988.

[9] Albert Visser. Interpretability Logic. To appear in: *Proceedings of the 1988 Heyting Colloquium*, Plenum Press, Boston, 1990.

The ITLI Prepublication Series

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
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
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
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