BI-UNARY INTERPRETABILITY LOGIC

Maarten de Rijke

ITLI Prepublication Series
X-90-12

University of Amsterdam
1986
86-01 The Institute of Language, Logic and Information
86-02 Peter van Emde Boas
86-03 Johan van Benthen
86-04 Reinhard Muskens
86-05 Kenneth A. Bowen, Dick de Jongh
86-06 Johan van Benthen
1987
87-01 Jeroen Groenendijk, Martin Stokhof
87-02 Renate Bartsch
87-03 Jan Willem Klof, Roel de Vrijer
87-04 Johan van Benthen
87-05 Víctor Sánchez Valencia
87-06 Eleonore Oversteegen
87-07 Johan van Benthen
87-08 Renate Bartsch
87-09 Herman Hendriks
1988
LP-88-01 Michel van Lambalgen Logic, Semantics and Philosophy of Language: Algorithmic Information Theory
LP-88-02 Yde Venema
LP-88-03 A.S. Troelstra
LP-88-04 Reinhard Muskens
LP-88-05 Johan van Benthen
LP-88-06 Dick de Jongh, Mark H. Overmars, Peter van Emde Boas
LP-88-07 Theo M.V. Jansen
LP-88-08 Annette Kleppe
ML-88-01 Jaap van Oosten
ML-88-02 M.D.G. Swaen
ML-88-03 Frank Veltman
ML-88-04 A.S. Troelstra
ML-88-05 A.S. Troelstra
CT-88-01 Ming Li, Paul M.B. Vitanyi Computability and Complexity Theory: Two Decades of Applied Kolmogorov Complexity
CT-88-02 Michel H.M. Smid
CT-88-03 Michel H.M. Smid, Mark H. Overmars, Peter van Emde Boas
CT-88-04 Dick de Jongh, Lex Hendriks
CT-88-05 Peter van Emde Boas
CT-88-06 Michel H.M. Smid
CT-88-07 Johan van Benthen
CT-88-08 Michel H.M. Smid, Mark H. Overmars
ML-88-01 Dick de Jongh, Albert Visser
ML-88-02 Roel de Vrijer
ML-88-03 Dick de Jongh, Franco Montagna
ML-88-04 Dick de Jongh, Marc Jumelet, Franco Montagna
ML-88-05 Rinke Verbrugge
ML-88-06 Michel van Lambalgen
ML-88-07 Dirk Roorda
ML-88-08 Alessandra Carbone
ML-88-09 Zhisheng Huang
ML-88-10 Edith Spaan, Leen Torenvliet, Peter van Emde Boas
ML-88-11 Sergei Danneheuvel
X-88-01 Marc Jumelet
1989
LP-89-01 Johan van Benthen Logic, Semantics and Philosophy of Language: The Fine-Structure of Categorial Semantics
LP-89-02 Jeroen Groenendijk, Martin Stokhof
LP-89-03 Yde Venema
LP-89-04 Johan van Benthen
LP-89-05 Johan van Benthen
LP-89-06 Andrej Prijatelj
LP-89-07 Heinrich Wansing
LP-89-08 Zhisheng Huang
LP-89-09 Zhisheng Huang
LP-89-10 Dick de Jongh, Albert Visser
ML-89-01 Dick de Jongh, Albert Visser
ML-89-02 Roel de Vrijer
ML-89-03 Dick de Jongh, Franco Montagna
ML-89-04 Dick de Jongh, Marc Jumelet, Franco Montagna
ML-89-05 Rinke Verbrugge
ML-89-06 Michel van Lambalgen
ML-89-07 Dirk Roorda
ML-89-08 Alessandra Carbone
ML-89-09 Zhisheng Huang
CT-89-01 Michel H.M. Smid
CT-89-02 Peter van Emde Boas
CT-89-03 Ming Li, Herman Nofzgile, Leen Torenvliet, Peter van Emde Boas
CT-89-04 Heman Nofzgile, Leen Torenvliet
CT-89-05 Piet ter Hertel, Michel H.M. Smid, Leen Torenvliet, Willem G. Vee
CT-89-06 H.W. Lenstra, Jr.
CT-89-07 Ming Li, Paul M.B. Vitanyi
CT-89-08 Harry Buhrman, Steven Homer
CT-89-09 Harry Buhrman, Edith Spaan, Leen Torenvliet
CT-89-10 Sieger van Danneheuvel
CT-89-11 Zhisheng Huang, Sieger van Danneheuvel Towards Functional Classification of Recursive Query Processing
X-89-01 Marianne Kalabek
X-89-02 G. Wagemakers
X-89-03 A.S. Troelstra
X-89-04 Jeroen Groenendijk, Martin Stokhof
X-89-05 Maarten de Rijke
X-89-06 Peter van Emde Boas
1990
SEE INSIDE BACK COVER
BI-UNARY INTERPRETABILITY LOGIC

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

Received August 1990

Research supported by the Netherlands Organization for Scientific Research (NWO)
Bi-Unary Interpretability Logic

Maarten de Rijke*

Department of Mathematics and Computer Science
University of Amsterdam

June 1990

1 Introduction

In recent years several modal systems have been introduced to study the relation of relative interpretability between arithmetical theories. The interpretability principles of several important classes of arithmetical theories have been axiomatised. In [6] the system ILP is shown to be the interpretability logic of all $\Sigma^0_1$-sound finitely axiomatised sequential theories that extend $\mathbf{I} \Delta_0 + \mathbf{SupExp}$; in [1] it is shown that ILM is the interpretability logic of $\mathbf{PA}$. Montagna and Hájek [2] show that ILM is also the logic of $\Pi^0_1$-conservativity of all $\Sigma^0_1$-sound extensions of $\Sigma_1$. (As is well-known, in the case of $\mathbf{PA}$ the two relations of relative interpretability and of $\Pi^0_1$-conservativity coincide).

Given the above results it is only natural to consider a modal logic with two binary modal operators, one of which is to be interpreted arithmetically as the relation of $\Pi^0_1$-conservativity between extensions of some given finitely axiomatised sequential extension $T$ of $\Sigma_1$, while the other operator is to be interpreted as relative interpretability over the same theory $T$. Such a system, called ILM/P, has been introduced by Dick de Jongh and Albert Visser, and is conjectured to be the logic of relative interpretability and $\Pi^0_1$-conservativity of all $\Sigma^0_1$-sound finitely axiomatised sequential extensions of $\Sigma_1$. Both the modal and arithmetical completeness of ILM/P are still open.

Interpretability may also be viewed as a unary predicate over extensions of a fixed theory $T$. The modal analysis of the interpretability predicate has been undertaken in [3], using, of course, a unary modal operator. In this note we axiomatize the bi-unary subsystem of ILM/P. That is, we introduce two unary operators $\mathbf{I}_M$, $\mathbf{I}_P$ with the following interpretations: $\mathbf{I}_MA$ stands for '$T + A$ is a $\Pi^0_1$-conservative extension of $T$', and $\mathbf{I}_PA$ stands for '$T + A$ is interpretable in

*Research supported by the Netherlands Organisation for Scientific Research (NWO).
$T'$, and we axiomatize all formulas $A$ in the language with only $\Box$, $I_M$, $I_P$ that are provable in $ILM/P$.

2 Axioms and models

The provability logic $L$ is propositional logic plus the axiom schemas $\Box (A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$, $\Box A \rightarrow \Box \Box A$ and $\Box (\Box A \rightarrow A) \rightarrow \Box A$, and the rules Modus Ponens ($\vdash A, \vdash A \rightarrow B \therefore \vdash B$) and Necessitation ($\vdash A \Rightarrow \vdash \Box A$). We use $L(\Box)$ to denote the language of $L$. $L(\Box)$ is extended with a binary operator $\triangleright$ to obtain the language $L(\Box, \triangleright)$ of binary interpretability logic. The binary interpretability logic $IL$ is obtained from $L$ by adding the axioms

$\begin{align*}
& (J1) \quad \Box (A \rightarrow B) \rightarrow A \triangleright B \\
& (J2) \quad (A \triangleright B) \land (B \triangleright C) \rightarrow (A \triangleright C) \\
& (J3) \quad (A \triangleright C) \land (B \triangleright C) \rightarrow (A \lor B) \triangleright C
\end{align*}$

where $\Diamond \equiv \neg \Box \rightarrow$. $ILM$ is $IL + M$ and $ILP$ is $IL + P$, where $M \equiv A \triangleright B \rightarrow A \land \Box C \triangleright B \land \Box C$ and $P \equiv A \triangleright B \rightarrow \Box (A \triangleright B)$.

The system $ILM/P$ is defined in a language $L(\Box, \triangleright_M, \triangleright_P)$ which contains the operator $\Box$ as well as two binary interpretability operators: $\triangleright_M$ and $\triangleright_P$. For the operator $\triangleright_M$ we assume the axioms $J1$–$J5$ and $M$; for the operator $\triangleright_P$ we assume the axioms $J1$–$J5$ and $P$. In addition there is one mixed axiom:

$A \triangleright_M B \rightarrow A \land (C \triangleright_P D) \triangleright_M B \land (C \triangleright_P D)$.

Define in $L(\Box, \triangleright)$ the unary interpretability operator ‘$\mathbf{T}$’ by $\mathbf{I}A := \mathbf{T} \triangleright A$, and let $L(\Box, \mathbf{I})$ extend $L(\Box)$ with $\mathbf{I}$. The unary interpretability logic $il$ is obtained from $L$ by adding the axioms

$\begin{align*}
& (I1) \quad \Box \mathbf{I} \mathbf{I} \\
& (I2) \quad \Box (A \rightarrow B) \rightarrow (\mathbf{I} A \rightarrow \mathbf{I} B) \\
& (I3) \quad \mathbf{I} (A \lor \Box A) \rightarrow \mathbf{I} A \\
& (I4) \quad \mathbf{I} A \land \mathbf{I} \mathbf{I} \rightarrow \mathbf{I} A.
\end{align*}$

We use $ilm$ to denote $il+m$ and $ilp$ to denote $il+p$, where $m \equiv \mathbf{I} A \rightarrow \mathbf{I} (A \land \Box)\mathbf{I}$ and $p \equiv \mathbf{I} A \rightarrow \Box \mathbf{I} A$.

In $L(\Box, \triangleright_M, \triangleright_P)$ we define the unary interpretability operators $I_M$ and $I_P$ by $I_M A := \mathbf{T} \triangleright_M A$ and $I_P A := \mathbf{T} \triangleright_P A$ respectively. It is sometimes convenient to assume that the unary system $ilm$ is defined in the language $L(\Box, I_M)$ with $\Box$ and $I_M$ as the only modal operators, and similarly for $ilp$ and $L(\Box, I_P)$. The system $ilm/p$ is defined in $L(\Box, I_M, I_P)$ as follows; it contains the axioms $I1$–$I4$ and $m$ for the operator $I_M$, and the axioms $I1$–$I4$ and $p$ for the operator $I_P$; it has no mixed axioms. (Note that $ilp \vdash m$, so in $ilm/p$ we also have axiom $m$ for the operator $I_P$.)

Recall that an $L$-frame is a pair $(W, R)$ with $R \subseteq W^2$ transitive and conversely well-founded, and that an $L$-model is given by an $L$-frame $F$ together
with a forcing relation \( \models \) that satisfies the usual clauses for \( \neg \) and \( \wedge \), while 
\( u \models A \) if and only if \( \forall v (uRv \Rightarrow v \models A) \). A (Veltman-) frame for \( IL \) is a triple \((W, R, S)\), where \((W, R)\) 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 \( \models \) that satisfies the above clauses for \( \neg \), \( \wedge \) and \( \Box \), while

\[
u \models A \rightarrow B \Leftrightarrow \forall v (uRv \land v \models A \Rightarrow \exists w (vS_w \land w \models B)).\]

An \( ILP \)-model is an \( IL \)-model that satisfies the extra condition: if \( wRw'RuS_wv \) then \( uS_{w'}v \). An \( ILM \)-model is an \( IL \)-model satisfying the extra condition: if \( uS_wvRz \) then \( uRz \).

An \( ILM/P \)-frame is a tuple \((W, R, S^M, S^P)\) such that \((W, R, S^M)\) is an \( ILM \)-frame, and \((W, R, S^P)\) is an \( ILP \)-frame, while the following extra condition connecting \( S^M \) and \( S^P \) holds:

\[
\forall xyzuv (xRys^M_xzRuS^P_yv \rightarrow uS^P_z).\]

An \( ILM/P \)-model is a tuple \((W, R, S^M, S^P, \models)\) such that \((W, R, S^M, S^P)\) is an \( ILM/P \)-frame, and such that the semantics of the operator \( \triangleright_M \) is based on the relation \( S^M \), while the semantics of the operator \( \triangleright_P \) is based on the relation \( S^P \).

The truth definition for \( I_K \) (\( K \in \{M, P\} \)) follows from the above definitions:

\[
z \models I_K A \text{ if and only if } \forall y (xRy \rightarrow \exists z (yS^K_z \land z \models A)).\]

## 3 Preliminaries

In this Section we introduce the tools needed to prove the modal completeness of \( ilm/p \). We start with some definitions.

**Definition 3.1** Let \( K \in \{M, P\} \), and let \( \Gamma, \Delta \) be two maximal \( ilm/p \)-consistent sets.

1. \( \Delta \) is called a \textit{successor} of \( \Gamma \) (\( \Gamma \prec \Delta \)) if
   
   (a) \( \Delta \), \( \Box A \in \Delta \) for each \( \Box A \in \Gamma \)
   
   (b) \( \Box A \in \Delta \) for some \( \Box A \not\in \Gamma \).

2. \( \Delta \) is called an \((I_K, C)\)-\textit{critical successor} of \( \Gamma \) if
(a) $\Gamma \prec \Delta$
(b) $I_K C \notin \Gamma$
(c) $\neg C, \square \neg C \in \Delta$.

Note that if $\Delta$ is a successor $\Gamma$ then it is both an $(I_M, \bot)$-critical and an $(I_P, \bot)$-critical successor of $\Gamma$.

**Proposition 3.2** Let $\Gamma$ be a maximal ilm/p-consistent set such that $\square C \in \Gamma$. Then there is a maximal ilm/p-consistent successor $\Delta$ of $\Gamma$ with $C, \square \neg C \in \Delta$.

**Proof.** Well-known (or cf. [4]). QED.

**Proposition 3.3** Let $K \in \{ M, P \}$, and let $\Gamma$ be a maximal ilm/p-consistent set such that $\neg I_K C \in \Gamma$. The there exists a maximal ilm/p-consistent $(I_K, C)$-critical successor $\Delta$ of $\Gamma$ with $\square \bot \in \Delta$.

**Proof.** Cf. [3, Proposition 2.4]. QED.

**Proposition 3.4** Let $K \in \{ M, P \}$, and let $I_K C \in \Gamma$, where $\Gamma$ is a maximal ilm/p-consistent set. If there exists a maximal ilm/p-consistent $(I_K, E)$-critical successor $\Delta$ of $\Gamma$, then there exists a maximal ilm/p-consistent $(I_K, E)$-critical successor $\Delta'$ of $\Gamma$ such that $C, \square \bot \in \Delta'$.

**Proof.** By axiom $m$, $I_K C$ implies $I_K (C \wedge \square \bot)$. By [3, Proposition 2.5] the result follows. QED.

Here is one more definition:

**Definition 3.5** A set of formulas $\Phi$ is called 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 $\neg B \in \Phi$

It is clear that every formula is contained in a finite adequate set.

4 The main theorem

Given some maximal ilm/p-consistent set $\Gamma$ and a finite adequate set $\Phi$, we define the structure $(W_\Gamma, R, S^M, S^P)$, which consists of pairs $(\Delta, \tau)$, where $\Delta$ is a maximal ilm/p-consistent set needed to handle the truth definition for formulas in $\Gamma$, and $\tau$ is a sequence of pairs we use to index the pairs we put into $W_\Gamma$.

For the time being, we fix a maximal ilm/p-consistent set $\Gamma$ and a finite adequate set $\Phi$. We use $\bar{w}, \bar{v}, \ldots$ to denote pairs $(\Delta, \tau)$. If $\bar{w} = (\Delta, \tau)$, then $(\bar{w})_0 = \Delta, (\bar{w})_1 = \tau$. We write $\sigma \subseteq \tau$ for $\sigma$ is an initial segment of $\tau$, and $\sigma \subset \tau$ if $\sigma$ is a proper initial segment of $\tau$. Finally, $\sigma \tau$ denotes the concatenation of $\sigma$ and $\tau$.  

4
Definition 4.1 Define $W_T$ to be a minimal set of pairs $(\Delta, \tau)$ such that

1. $(\Gamma, (\langle \rangle)) \in W_T$;
2. if $(\Delta, \tau) \in W_T$, $\square B \in \Delta \cap \Phi$, and if there exists a successor $\Delta'$ of $\Delta$ with $B$, $\square \neg B \in \Delta'$, then $(\Delta', \tau \neg \langle \langle \square B, \perp \rangle \rangle) \in W_T$ for one such $\Delta'$;
3. if $(\Delta, \tau) \in W_T$, $\neg I_M B \in \Delta \cap \Phi$, and if there exists an $(\langle I_M, B \rangle)$-critical successor $\Delta'$ of $\Delta$ with $\square \perp \in \Delta'$, then $(\Delta', \tau \neg \langle \langle \neg I_M B, B \rangle \rangle) \in W_T$ for one such $\Delta'$;
4. if $(\Delta, \tau) \in W_T$, $\neg I_P B \in \Delta \cap \Phi$, and if there exists an $(\langle I_P, B \rangle)$-critical successor $\Delta'$ of $\Delta$ with $\square \perp \in \Delta'$, then $(\Delta', \tau \neg \langle \langle \neg I_P B, B \rangle \rangle) \in W_T$ for one such $\Delta'$;
5. if $(\Delta, \tau) \in W_T$, $I_M B \in \Delta \cap \Phi$, $C \in \Phi$, and if there exists an $(\langle I_M, C \rangle)$-critical successor $\Delta'$ of $\Delta$ with $B$, $\square \perp \in \Delta'$, then $(\Delta', \tau \langle \langle I_M B, C \rangle \rangle) \in W_T$ for one such $\Delta'$;
6. if $(\Delta, \tau) \in W_T$, $I_P B \in \Delta \cap \Phi$, $C \in \Phi$, and if there exists an $(\langle I_P, C \rangle)$-critical successor $\Delta'$ of $\Delta$ with $B$, $\square \perp \in \Delta'$, then $(\Delta', \tau \langle \langle I_P B, C \rangle \rangle) \in W_T$ for one such $\Delta'$.

Define $R$ on $W_T$ by putting $\bar{w}R\bar{v}$ if $(\bar{w})_1 \subseteq (\bar{v})_1$.
Define $S^M$ on $W_T$ by putting $\bar{w}S^M_{\bar{v}} \bar{u}$ iff for some $B, B', C, C', \sigma$ and $\sigma'$:

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

and either $(\bar{v})_1 \subseteq (\bar{u})_1$, or $B$ is not of the form $I_M D$ or $I_M D'$, and then $B' \equiv I_M D'$ and $C' \equiv \perp$ for some $D'$, or $B$ is of the form $I_M D$ or $I_M D'$, and then $C' \equiv C$ and $B' \equiv I_M D'$ for some $D'$.

Define $S^P$ on $W_T$ by putting $\bar{w}S^P_{\bar{u}} \bar{v}$ iff for some $B, B', C, C', \tau, \tau'$ and $\sigma$:

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

and either $(\bar{v})_1 \subseteq (\bar{u})_1$, or $B$ is not of the form $I_P D$ or $I_P D'$, and then $B' \equiv I_P D'$ and $C' \equiv \perp$ for a $D'$, or $B$ is of the form $I_P D$ or $I_P D'$, and then $C' \equiv C$ and $B' \equiv I_P D'$ for some $D'$.

Proposition 4.2 1. $(W_T, R, S^M, S^P)$ is finite.
2. If $\bar{w} \in W_T$, and $(\bar{w})_1 = \tau \langle \langle \neg I_K B, C \rangle \rangle \langle \sigma \rangle$, where $K \in \{ M, P \}$, then $\bar{w}$ is an $R$-endpoint, $\square \perp \in (\bar{w})_0$, and $\sigma = (\langle \rangle)$.
3. If $\bar{u} \in W_T$, $(\bar{u})_1 = \tau \langle \langle \perp B, \perp \rangle \rangle$, and if we have $\bar{w}S^M_{\bar{u}} \bar{v}$ or $\bar{u}S^P_{\bar{u}} \bar{v}$, then $\bar{w}R\bar{v}R\bar{u}$.
4. If $(\bar{w})_1 = (\bar{v})_1$ then $\bar{w} = \bar{v}$.
5. If $\bar{w}R\bar{u}$ then $\bar{w}_0 \prec (\bar{v})_0$.
6. $(W_T, R)$ is a tree.
7. $(W_T, R, S^M)$ is an ILM-frame.
8. $(W_T, R, S^P)$ is an ILP-frame.
9. $(W_T, R, S^M, S^P)$ is an ILM/P-frame.
Proof. Left to the reader. QED.

Theorem 4.3 Let $A \in \mathcal{L}(\Box, I_M, I_P)$. Then $\text{ilmp} \vdash A$ iff for all finite ILM/P-models $\mathcal{M}$ we have $\mathcal{M} \models A$.

Proof. We only prove completeness. Assume $\text{ilmp} \not\vdash A$. Let $\Gamma$ be a maximal $\text{ilmp}$-consistent set with $\neg A \in \Gamma$, and let $\Phi$ be a finite adequate set with $\neg A \in \Phi$. Construct $(W_T, R, S^M, S^P)$ as in 4.1. We complete the proof by putting $\overline{w} \models p$ iff $p \in (\overline{w})_0$, and by proving that for all $F \in \Phi$ and $\overline{w} \in W_T$, we have $\overline{w} \models F$ iff $F \in (\overline{w})_0$. The proof is by induction on $F$. We only consider the cases $F \equiv \Box C, I_M D$ and $I_P D$.

If $F \equiv \Box C \in (\overline{w})_0$, then we have to show that $\exists \overline{v} (\overline{v} R \overline{u} \land B \in (\overline{v})_0)$. Now, by 3.2 there exists a successor $\Delta$ of $(\overline{w})_0$ with $B, \Box \bot \in \Delta$. We may assume that $\overline{u} := \langle \Delta, (\overline{w})_1 \sim \langle (\Box B, \bot) \rangle \rangle \in W_T$. Obviously, $\overline{u} R \overline{u}$ and $B \in (\overline{u})_0$, as required.

The case $F \equiv \Box C \notin (\overline{w})_0$ is trivial.

Assume that $I_M D \in (\overline{w})_0$. We have to show that $\forall \overline{v} (\overline{v} R \overline{u} \rightarrow \exists \overline{w} (\overline{w} S^M \overline{u} \land D \in (\overline{u})_0))$. Assume that $\overline{w} R \overline{u}$; then for some $B, C$ and $\sigma$, $(\overline{w})_1 \sim \langle (B, C) \rangle \sim \sigma$. If $B$ is not of the form $\langle \neg \rangle I_M B'$, then we consider $(\overline{w})_0$ to be an $(I_M, \bot)$-critical successor of $(\overline{w})_0$. By 3.4 there exists an $(I_M, \bot)$-critical successor $\Delta$ of $(\overline{w})_0$ with $D, \Box \bot \in \Delta$. Put $\overline{u} := \langle \Delta, (\overline{w})_1 \sim \langle (I_M D, \bot) \rangle \rangle$. We may assume that $\overline{u} \in W_T$. It is clear that $\overline{w} S^M \overline{u}$ and $D \in (\overline{u})_0$, as required. Next we suppose that $B$ is of the form $\langle \neg \rangle I_M B'$. Then $(\overline{w})_0$ is an $(I_M, C)$-critical successor of $(\overline{w})_0$. By 3.4 there exists an $(I_M, C)$-critical successor $\Delta$ of $(\overline{w})_0$ with $D, \Box \bot \in \Delta$. Put $\overline{u} := \langle \Delta, (\overline{w})_1 \sim \langle (I_M D, C) \rangle \rangle$. Then we may assume that $\overline{u} \in W_T$. Moreover, we have $\overline{w} S^M \overline{u}$ and $D \in (\overline{u})_0$, as required.

Assume that $I_P D \notin (\overline{w})_0$. Then $\neg I_M D \in (\overline{w})_0$. We have to prove that $\exists \overline{v} (\overline{v} R \overline{u} \land \forall \overline{w} (\overline{w} S^P \overline{u} \rightarrow D \notin (\overline{u})_0))$. Now, by 3.3 there exists an $(I_P, D)$-critical successor $\Delta$ of $(\overline{w})_0$ with $\Box \bot \in \Delta$. We may assume that $\overline{u} := \langle \Delta, (\overline{w})_1 \sim \langle (I_P D, \bot) \rangle \rangle \in W_T$. Now suppose that for some $\overline{u} \in W_T$. By definition $(\overline{w})_1 \sim \langle (B', C') \rangle \sim \sigma'$, for some $B'$, $C'$ and $\sigma'$. Since $\Box \bot \in (\overline{w})_0$, we can not have $\overline{v} R \overline{u}$. Hence, we have either $\overline{u} = \overline{v}$ and then $D \notin (\overline{u})_0$, or $C' = D$ and $B' \equiv I_M D'$ for some $D'$. But then $(\overline{w})_0$ must be an $(I_P, D')$-critical successor of $(\overline{w})_0$—and so $D \notin (\overline{u})_0$.

Assume that $I_P D \in (\overline{w})_0$. We have to show that $\forall \overline{v} (\overline{v} R \overline{u} \rightarrow \exists \overline{w} (\overline{w} S^P \overline{u} \land D \in (\overline{u})_0))$. So assume that $\overline{u} R \overline{v}$. Since $(W_T, R)$ is a tree, we can find a unique immediate $R$-predecessor $\overline{w}'$ of $\overline{v}$. By axiom $p$ (for $I_P$) we must have $I_P D \in (\overline{w})_0$. Hence, by axiom $m$ for $I_P$, also $I_P (D \land \Box \bot) \in (\overline{w})_0$. By construction $(\overline{w})_1 \sim \langle (B, C) \rangle$ for some $B$ and $C$. If $B$ is not of the form $\langle \neg \rangle I_P B'$, then we consider $(\overline{w})_0$ to be an $(I_P, \bot)$-critical successor of $(\overline{w})_0$. By 3.4 there exists an $(I_P, \bot)$-critical successor $\Delta$ of $(\overline{w})_0$ with $D, \Box \bot \in \Delta$. We may assume that $\overline{u} := \langle \Delta, (\overline{w})_1 \sim \langle (I_P D, \bot) \rangle \rangle \in W_T$. Moreover it is clear that $\overline{w} S^P \overline{u}$ and $D \in (\overline{u})_0$, as required. If, on the other hand, $B$ is of the form $\langle \neg \rangle I_P B'$, then $(\overline{w})_0$ is an $(I_P, C)$-critical successor of $(\overline{w})_0$. By 3.4 there exists an $(I_P, C)$-critical successor $\Delta$ of $(\overline{w})_0$ with $D, \Box \bot \in \Delta$. As before we may
assume that $\bar{u} := (\Delta, (\bar{w}), (\mathfrak{I}_PD, C)) \in W_T$. Moreover, we have $\psi \mathfrak{S}_w \bar{u}$ and $D \in \mathfrak{K}_0$, as required.

The last case we have to consider is the case that $\mathfrak{I}_PD \notin (\bar{w}),0$. But this case is entirely analogous to the case $\mathfrak{I}_M D \notin (\bar{w}),0$. QED.

**Proposition 4.4** Let $A \in \mathcal{L}(\mathfrak{M}, \mathfrak{I}_M, \mathfrak{I}_P)$. Then $\mathfrak{u}m/p \vdash A$ iff $\mathfrak{I}M/P \vdash A$.

**Proof.** If $\mathfrak{u}m/p \vdash A$ then, by a simple induction on derivations, $\mathfrak{I}M/P \vdash A$. If $\mathfrak{u}m/p \not\vdash A$ then by 4.3 there is a finite $\mathfrak{I}M/P$-model $\mathcal{M}$ with $\mathcal{M} \not\models A$. By the soundness of $\mathfrak{I}M/P$ w.r.t. $\mathfrak{I}M/P$-models it follows that $\mathfrak{I}M/P \not\models A$. QED.

**Proposition 4.5** Let $A \in \mathcal{L}(\mathfrak{M}, \mathfrak{I}_M)$. Then $\mathfrak{u}m/p \vdash A$ iff $\mathfrak{u}m \vdash A$ iff $\mathfrak{I}M \vdash A$.

**Proof.** The second equivalence is [3, Proposition 2.15]. If $\mathfrak{u}m \vdash A$ then obviously $\mathfrak{u}m/p \vdash A$. And if $\mathfrak{u}m \not\vdash A$ then by [3, Theorem 2.14] there is an $\mathfrak{I}M$-model $\mathcal{M}$ with $\mathcal{M} \not\models A$. $\mathcal{M}$ may be turned into an $\mathfrak{I}M$-model $\mathcal{M}'$ by defining $\psi^*_Z \equiv \exists \mathfrak{R}_z \mathfrak{R}_x$. Obviously, $\mathcal{M}' \not\models A$. So by 4.3 $\mathfrak{u}m/p \not\vdash A$. QED.

**Proposition 4.6** Let $A \in \mathcal{L}(\mathfrak{M}, \mathfrak{I}_P)$. Then $\mathfrak{u}m/p \vdash A$ iff $\mathfrak{u}p \vdash A$ iff $\mathfrak{I}L \vdash A$.

**Proof.** Similar to the proof of 4.5—using [3, Proposition 2.25 and Theorem 2.23]. QED.

Fix $T$ to be a $\Sigma_1^0$-sound finitely axiomatised sequential extension of $\Pi_1$, and define the arithmetical interpretation $(\cdot)^*$ of $\mathcal{L}(\mathfrak{M}, \mathfrak{I}_M, \mathfrak{I}_P)$ into the language of $T$ as usual for proposition letters, Boolean connectives and $\mathfrak{K}$, while

$(\mathfrak{I}_P A)^* := 'T^* + A^*$ is interpretable in $T^*$

$(\mathfrak{I}_M A)^* := 'for all $\Pi_1^0$-sentences $\varphi$, if $\varphi$ is provable in $T^* + A^*$, then $\varphi$ is provable in $T^*$."

**Proposition 4.7** 1. Let $A \in \mathcal{L}(\mathfrak{M}, \mathfrak{I}_M)$. Then $\mathfrak{u}m/p \vdash A$ iff for all $(\cdot)^*$, $T^* \vdash A^*$.

2. Let $A \in \mathcal{L}(\mathfrak{K}, \mathfrak{I}_P)$. Then $\mathfrak{u}m/p \vdash A$ iff for all $(\cdot)^*$, $T^* \vdash A^*$.

**Proof.** To prove (1) use 4.5 and the fact that by [5, Theorem 10.1], $\mathfrak{I}M \vdash A$ iff for all interpretations $(\cdot)^*$ of $\mathcal{L}(\mathfrak{M}, \mathfrak{I}_M)$ into the language of $T^*$, $T^* \vdash A^*$. To prove (2) use 4.6 and the fact that by [6, Theorem 8.2], $\mathfrak{I}L \vdash A$ iff for all interpretations $(\cdot)^*$ of $\mathcal{L}(\mathfrak{K}, \mathfrak{I}_P)$ into the language of $T^*$, $T^* \vdash A^*$. QED.

According to Propositions 4.4 and 4.7 what $\mathfrak{I}M/P$ says about unary interpretability and unary $\Pi_1^0$-conservativity considered separately is precisely what it should say about these predicates. This leads additional support to the conjecture that $\mathfrak{I}M/P$ is the logic of the relations of relative interpretability and $\Pi_1^0$-conservativity (taken together) of all $\Sigma_1^0$-sound finitely axiomatized sequential extensions of $\Pi_1$. 

7
References


The ITLI Prepublication Series

1990

Logic, Semantics and Philosophy of Language

LP-90-01 Jaap van der Does
LP-90-02 Jeroen Groenendijk, Martin Stokhof
LP-90-03 Renate Bartsch
LP-90-04 Aarne Ranta
LP-90-05 Patrick Blackburn
LP-90-06 Gennaro Chierchia
LP-90-07 Gennaro Chierchia
LP-90-08 Herman Hendriks
LP-90-09 Paul Dekker
LP-90-10 Theo M.V. Janssen
LP-90-11 Johan van Benthem

Mathematical Logic and Foundations

ML-90-01 Harold Schellinx
ML-90-02 Jaap van Oosten
ML-90-03 Yde Venema
ML-90-04 Maarten de Rijke
ML-90-05 Domenico Zambrulla
ML-90-06 Jaap van Oosten
ML-90-07 Maarten de Rijke

Computation and Complexity Theory

CT-90-01 John Tromp, Peter van Emde Boas
CT-90-02 Sieger van Denneheuvel
Gerard R. Renardel de Lavalette

CT-90-03 Ricard Gavalda, Leen Torenvliet
Osamu Watanabe, José L. Balcázar
CT-90-04 Harry Buhrman, Leen Torenvliet

Other Prepublications

X-90-01 A.S. Troelstra

X-90-02 Maarten de Rijke
X-90-03 L.D. Beklemishev
X-90-04
X-90-05 Valentin Shehtman
X-90-06 Valentin Goranko, Solomon Passy
X-90-07 V. Yu. Shavrukov
X-90-08 L.D. Beklemishev

X-90-09 V. Yu. Shavrukov
X-90-10 Sieger van Denneheuvel
Peter van Emde Boas
X-90-11 Alessandra Carbone
X-90-12 Maarten de Rijke

A Generalized Quantifier Logic for Naked Infinitives
Dynamic Montague Grammar
Concept Formation and Concept Composition
Intuitionistic Categorial Grammar
Nominal Tense Logic
The Variability of Impersonal Subjects
Anaphora and Dynamic Logic
Flexible Montague Grammar
The Scope of Negation in Discourse, towards a flexible dynamic Montague grammar
Models for Discourse Markers
General Dynamics

Isomorphisms and Non-Isomorphisms of Graph Models
A Semantical Proof of De Jongh’s Theorem
Relational Games
Unary Interpretability Logic
Sequences with Simple Initial Segments
Extension of Lifschitz’ Realizability to Higher Order Arithmetic, and a Solution to a Problem of F. Richman
A Note on the Interpretability Logic of Finitely Axiomatized Theories

Associative Storage Modification Machines
A Normal Form for PCSJ Expressions

Generalized Kolmogorov Complexity
in Relativized Separations
Bounded Reductions

Remarks on Intuitionism and the Philosophy of Mathematics,
Revised Version
Some Chapters on Interpretability Logic
On the Complexity of Arithmetical Interpretations of Modal Formulae
Annual Report 1989
Derived Sets in Euclidean Spaces and Modal Logic
Using the Universal Modality: Gains and Questions
The Lindenbaum Fixed Point Algebra is Undecidable
Provability Logics for Natural Turing Progressions of Arithmetical Theories
On Rosser’s Provability Predicate
An Overview of the Rule Language RL/1

Provable Fixed points in $\mathcal{L}_3 + \Omega_2$, revised version
Bi-Unary Interpretability Logic