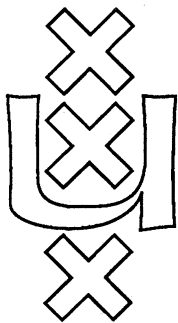


**Institute for Language, Logic and Information**

**ON THE COMPLEXITY OF ARITHMETICAL  
INTERPRETATIONS OF MODAL FORMULAE**

L.D. Beklemishev

ITLI Prepublication Series  
X-90-03



**University of Amsterdam**

## The ITLI Prepublication Series

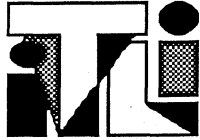
1986

- 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 Categorical 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,  
 86-06 Johan van Benthem Logical Syntax Forward looking Operators  
 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 Victor 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 Categorical Grammar and Type Theory  
 87-08 Renate Bartsch The Construction of Properties under Perspectives  
 87-09 Herman Hendriks Type Change in Semantics: The Scope of Quantification and Coordination

1988

- LP-88-01 Michiel van Lambalgen *Logic, Semantics and Philosophy of Language:* Algorithmic Information Theory  
 LP-88-02 Yde Venema Expressiveness and Completeness of an Interval Tense Logic  
 LP-88-03 Year Report 1987  
 LP-88-04 Reinhard Muskens Going partial in Montague Grammar  
 LP-88-05 Johan van Benthem Logical Constants across Varying Types  
 LP-88-06 Johan van Benthem Semantic Parallels in Natural Language and Computation  
 LP-88-07 Renate Bartsch Tenses, Aspects, and their Scopes in Discourse  
 LP-88-08 Jeroen Groenendijk, Martin Stokhof Context and Information in Dynamic Semantics  
 LP-88-09 Theo M.V. Janssen A mathematical model for the CAT framework of Eurotra  
 LP-88-10 Anneke Kleppe A Blissymbolics Translation Program  
 ML-88-01 Jaap van Oosten *Mathematical Logic and Foundations:* Lifschitz' Realizability  
 ML-88-02 M.D.G. Swaen The Arithmetical Fragment of Martin Löf's Type Theories with weak  $\Sigma$ -elimination  
 ML-88-03 Dick de Jongh, Frank Veltman Provability Logics for Relative Interpretability  
 ML-88-04 A.S. Troelstra On the Early History of Intuitionistic Logic  
 ML-88-05 A.S. Troelstra Remarks on Intuitionism and the Philosophy of Mathematics  
 CT-88-01 Ming Li, Paul M.B. Vitanyi *Computation and Complexity Theory:* Two Decades of Applied Kolmogorov Complexity  
 CT-88-02 Michiel H.M. Smid General Lower Bounds for the Partitioning of Range Trees  
 CT-88-03 Michiel H.M. Smid, Mark H. Overmars Maintaining Multiple Representations of  
 Leen Torenvliet, Peter van Emde Boas Dynamic Data Structures  
 CT-88-04 Dick de Jongh, Lex Hendriks Computations in Fragments of Intuitionistic Propositional Logic  
 Gerard R. Renardel de Lavalette  
 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 Multiple Representations of Dynamic Data Structures  
 Leen Torenvliet, Peter van Emde Boas  
 CT-88-09 Theo M.V. Janssen Towards a Universal Parsing Algorithm for Functional Grammar  
 CT-88-10 Edith Spaan, Leen Torenvliet, Peter van Emde Boas Nondeterminism, Fairness and a Fundamental Analogy  
 CT-88-11 Sieger van Denneheuvel, Peter van Emde Boas Towards implementing RL  
 X-88-01 Marc Jumelet *Other prepublications:* On Solovay's Completeness Theorem  
 1989 LP-89-01 Johan van Benthem *Logic, Semantics and Philosophy of Language:* The Fine-Structure of Categorical Semantics  
 LP-89-02 Jeroen Groenendijk, Martin Stokhof Dynamic Predicate Logic, towards a compositional,  
 non-representational semantics of discourse  
 LP-89-03 Yde Venema Two-dimensional Modal Logics for Relation Algebras and Temporal Logic of Intervals  
 LP-89-04 Johan van Benthem Language in Action  
 LP-89-05 Johan van Benthem Modal Logic as a Theory of Information  
 LP-89-06 Andreja Prijatelj Intensional Lambek Calculi: Theory and Application  
 LP-89-07 Heinrich Wansing The Adequacy Problem for Sequential Propositional Logic  
 LP-89-08 Víctor Sánchez Valencia Peirce's Propositional Logic: From Algebra to Graphs  
 LP-89-09 Zhisheng Huang Dependency of Belief in Distributed Systems  
 ML-89-01 Dick de Jongh, Albert Visser *Mathematical Logic and Foundations:* Explicit Fixed Points for Interpretability Logic  
 ML-89-02 Roel de Vrijer Extending the Lambda Calculus with Surjective Pairing is conservative  
 ML-89-03 Dick de Jongh, Franco Montagna Rosser Orderings and Free Variables  
 ML-89-04 Dick de Jongh, Marc Jumelet, Franco Montagna On the Proof of Solovay's Theorem  
 ML-89-05 Rineke Verbrugge  $\Sigma$ -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  $\text{ID}_0 + \Omega_1$   
 CT-89-01 Michiel H.M. Smid *Computation and Complexity Theory:* Dynamic Deferred Data Structures  
 CT-89-02 Peter van Emde Boas Machine Models and Simulations  
 CT-89-03 Ming Li, Herman Neuféglise, Leen Torenvliet, Peter van Emde Boas On Space Efficient Simulations  
 CT-89-04 Harry Buhrman, Leen Torenvliet A Comparison of Reductions on Nondeterministic Space  
 CT-89-05 Pieter H. Hartel, Michiel H.M. Smid A Parallel Functional Implementation of Range Queries  
 Leen Torenvliet, Willem G. Vree  
 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 Honest Reductions, Completeness and  
 Leen Torenvliet Nondeterministic 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 Towards Functional Classification of Recursive Query Processing  
 Peter van Emde Boas  
 X-89-01 Marianne Kalsbeek *Other Prepublications:* An Orey Sentence for Predicative Arithmetic  
 X-89-02 G. Wagemakers New Foundations: a Survey of Quine's Set Theory  
 X-89-03 A.S. Troelstra Index of the Heyting Nachlass  
 X-89-04 Jeroen Groenendijk, Martin Stokhof Dynamic Montague Grammar, a first sketch  
 X-89-05 Maarten de Rijke The Modal Theory of Inequality  
 X-89-06 Peter van Emde Boas Een Relationele Semantiek voor Conceptueel Modelleren: Het RL-project

1990 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 Muidersgracht 24  
1018TV Amsterdam

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

## **ON THE COMPLEXITY OF ARITHMETICAL INTERPRETATIONS OF MODAL FORMULAE**

L.D. Beklemishev  
Steklov Mathematical Institute  
Vavilova 42, 117966 Moscow GSP-1, U.S.S.R.



ON THE COMPLEXITY OF ARITHMETICAL  
INTERPRETATIONS OF MODAL FORMULAE

Beklemishev L.D.

Steklov Mathematical Institute,  
Moscow, November 1989.

It is a well-known fact that for any arithmetic sentence  $A$

$$A \in \Sigma_1^{\text{PA}} \Rightarrow \text{PA} \vdash A \rightarrow \text{Pr} \ulcorner A \urcorner.$$

Here  $\text{Pr}$  stands for Gödel's formula expressing provability in Peano Arithmetic  $\text{PA}$  and  $\Sigma_1^{\text{PA}}$  denotes the class of sentences  $\text{PA}$ -equivalent to those in  $\Sigma_1$ -form.

C.Kent [1] showed that the converse implication does not hold. Moreover, he found that for each natural number  $n$  there exists an arithmetic sentence  $A$  such that

$$\text{PA} \vdash A \rightarrow \text{Pr} \ulcorner A \urcorner \quad \text{and} \quad A \notin \Delta_n^{\text{PA}}.$$

D.Guaspari [2] rediscovered (a sharpened version of) this result applying his own techniques of partially conservative sentences. He also showed that arithmetically complex sentences implying their own provability cannot be constructed by some class of restricted means.

D.Guaspari posed a few problems generalizing one solved by Kent and himself, which are formulated in terms of provability interpretations of propositional modal logic.

DEFINITION. Let  $\mathcal{L}$  be the language consisting of propositional variables  $p, q, \dots$ ; boolean connectives  $\wedge, \vee, \rightarrow, \leftrightarrow, \neg$  and  $\perp$ ; modal operator  $\Box$ .

An *arithmetical interpretation*  $f$  is a mapping of  $\mathcal{L}$ -formulae to arithmetic sentences which commutes with boolean connectives and translates  $\Box$  as provability, i.e. for every modal formula  $\phi$

$$f(\Box\phi) = \text{Pr} \ulcorner f(\phi) \urcorner.$$

D.Guaspari noted that for any arithmetic sentence  $A$  obviously  $\text{PA} \vdash A \rightarrow \text{Pr} \ulcorner A \urcorner \Leftrightarrow$  there is a sentence  $B$  s.t.  $\text{PA} \vdash A \leftrightarrow B \wedge \text{Pr} \ulcorner B \urcorner$ .

So, the original question is equivalent to one, whether there exist non  $\Sigma_1^{\text{PA}}$  ( or even arbitrarily complex ) arithmetic interpretations of the modal formula  $p \wedge \Box p$ . D.Guaspari set the problem of cha-

racterizing those formulae of  $\mathcal{L}$ , which are  $\Sigma_1^{\text{PA}}$  under every arithmetical interpretation, and conjectured that they are exactly the formulae provably in the modal logic GL (called PRL in [3]) equivalent to disjunctions of boxed formulae, i.e. those of the form  $\Box\phi$ . An analogous conjecture was also made for the modal language enriched by witness comparison formulae.

These conjectures have been proved by A.Visser [4] and D.de Jongh & D.Pianigiani [5] respectively. The latter authors also extended the result to arbitrary theories containing arithmetic  $\text{I}\Sigma_1$ .

The aim of present paper is to characterize the formulae of  $\mathcal{L}$  having bounded arithmetical complexity, i.e. whose arithmetical interpretations are all  $\Delta_n^{\text{PA}}$  for some fixed  $n$ . It turns out, perhaps not very surprisingly, that such formulae are exactly those equivalent in GL to boolean combinations of boxed ones. In other words, each modal formula not equivalent to any such combination admits arbitrarily complex arithmetical interpretations. Thus, we may say that the "modal analog" of the arithmetical hierarchy with respect to PA collapses after the level of boolean combinations of  $\Sigma_1$ -sentences.

It is easily seen that the formula  $p\wedge\Box p$  is not equivalent to any boolean combination of boxed formulae, so our result provides us with one more proof of Kent's theorem. However, it will be clear after reading the sequel that, when applied to this particular formula, our construction essentially goes along the lines of that of Kreisel & Lévy (cf also [6], [7]).

Let  $B_n^{\text{PA}}$  denote the collection of all arithmetic sentences PA-equivalent to boolean combinations of those in  $\Sigma_n$ -form.

**THEOREM.**

Suppose  $\phi$  is a formula of  $\mathcal{L}$  such that for no boolean combination  $\psi$  of boxed formulae  $\text{GL} \vdash \phi \leftrightarrow \psi$ . Then for each  $n \geq 1$  there exists an arithmetical interpretation  $f$  such that

$$f(\phi) \in \Delta_{n+1}^{\text{PA}} \setminus B_n^{\text{PA}}.$$

In order to prove this theorem first of all we obtain some Kripke-style characterization of modal formulae GL-equivalent to boolean combinations of boxed formulae. We shall call the usual finite strictly partially ordered (not necessary treelike) Kripke models for GL just *models* (cf [3]). The expression  $\mathcal{K} \models \phi$  will

denote the fact that  $\phi$  is forced at the bottom node of the model  $\mathcal{K}$ .

DEFINITION. A modal formula  $\phi$  is called *stable* iff for all models  $\mathcal{K}$  and  $\mathcal{K}'$  with the same frame and the same forcing of propositional variables at all nodes except bottom ones

$$\mathcal{K} \Vdash \phi \Leftrightarrow \mathcal{K}' \Vdash \phi.$$

LEMMA 1. A modal formula  $\phi$  is GL-equivalent to a boolean combination of boxed formulae iff  $\phi$  is stable.

PROOF. It is trivial that boolean combinations of boxed formulae are stable. We check the converse implication.

Suppose  $\phi$  is stable. Clearly  $\phi$  is a boolean combination of boxed subformulae  $\Box\phi_1, \dots, \Box\phi_k$  and propositional variables  $p_1, \dots, p_m$  occurring outside any  $\Box$  in  $\phi$ . Let  $\phi^*$  denote the result of substituting in  $\phi$  the constant  $\perp$  for all the occurrences of variables outside any  $\Box$ . Clearly,  $\phi^*$  is a boolean combination of boxed formulae. We shall show that  $\text{GL} \vdash \phi \leftrightarrow \phi^*$ . By the completeness of GL with respect to models it is sufficient to show that  $\phi$  and  $\phi^*$  are forced exactly at the same models (cf [3]).

Let  $\mathcal{K}$  be an arbitrary model and let  $\mathcal{K}^*$  denote the model with the same frame as  $\mathcal{K}$  and the same forcing of propositional variables at all nodes except the bottom node, where (in  $\mathcal{K}^*$ ) no variable is forced. Clearly we have

$$\begin{aligned} \mathcal{K} \Vdash \Box\phi_i &\Leftrightarrow \mathcal{K}^* \Vdash \Box\phi_i \quad \text{and} \quad \mathcal{K}^* \Vdash p_j \Leftrightarrow \mathcal{K} \Vdash \perp, \text{ hence} \\ \mathcal{K} \Vdash \phi^* &\Leftrightarrow \mathcal{K}^* \Vdash \phi. \end{aligned}$$

By our assumption  $\phi$  is stable, hence

$$\mathcal{K} \Vdash \phi \Leftrightarrow \mathcal{K}^* \Vdash \phi.$$

It follows that

$$\mathcal{K} \Vdash \phi \Leftrightarrow \mathcal{K} \Vdash \phi^*.$$

Thus,

$$\text{GL} \vdash \phi \leftrightarrow \phi^*.$$

COROLLARY 1. The class of  $\mathcal{L}$ -formulae equivalent to boolean combinations of boxed ones is decidable.

Suppose now that  $\phi$  is a nonstable formula. Let  $\mathcal{K}_1$  and  $\mathcal{K}_2$  be a pair of models obtained by Lemma 1, such that  $\mathcal{K}_1 \Vdash \phi$  and  $\mathcal{K}_2 \Vdash \neg\phi$ . By identifying the corresponding nodes of  $\mathcal{K}_1$  and  $\mathcal{K}_2$  except the bottom ones and by adding a new bottom node below them we can construct a model  $\mathcal{K} = (\mathcal{K}, <, \Vdash)$ , such that

- (i)  $K = \{0, \dots, k\}$  for some  $k \in \mathbb{N}$ ;
- (ii) 0 is the bottom node of  $K$  and 1 and 2 are the only immediate successors of 0;
- (iii) 1 and 2 have the same successors;
- (iv) for all  $i, z, w \in K$ , if  $1 < z, w$  and  $z, w < i$  then either  $z < w$  or  $w < z$ ;
- (v)  $1 \Vdash \phi$  and  $2 \Vdash \neg \phi$ .

Condition (iv) will be satisfied if we beforehand choose  $K_1$  and  $K_2$  treelike .

Our next goal is to describe a Solovay-style [8] embedding of such models into arithmetic. First of all, let us try to explain it informally. Since we are going to define a very complex interpretation, say  $f$ , it is of no importance for us, whether we shall work within PA or within some finite extension of it of low complexity. The new axiom we will add is the statement, that some Solovay-type function has the limit  $\ell$  either in the node 1 or in 2. The words "Solovay-type" mean that for all  $i \in K \setminus \{0\}$  and all formulae  $\psi$  of  $\mathcal{L}$  we will have

$$PA \vdash \ell = i \rightarrow (i \Vdash \psi \leftrightarrow f(\psi)).$$

Thus, since  $1 \Vdash \phi$  and  $2 \Vdash \neg \phi$ , we shall obtain

$$PA \vdash \ell \in \{1, 2\} \rightarrow (\ell = 1 \leftrightarrow f(\phi)).$$

At this point all we need is to define our function in such a way, as it would be an impracticable task not only for  $PA + \ell \in \{1, 2\}$ , but also for any finite consistent extension of  $PA + \ell \in \{1, 2\}$  of low complexity, to distinguish whether  $\ell$  equals 1 or 2. And it is here that we apply the trick of Kreisel & Lévy. It is worth mentioning that such an idea would not work, had not our frame been fully symmetrical with respect to transpositions of nodes 1 and 2.

Now we turn to explicit definitions. Let for  $n \geq 1$   $Tr_n$  denote the standard  $\Delta_{n+1}$ -definition of truth for  $B_n$ -formulae, i.e. the standard arithmetic  $\Delta_{n+1}$ -formula such that

$$PA \vdash \forall u \in B_n \text{ Pr}( u \leftrightarrow \ulcorner Tr_n(u) \urcorner )$$

and

$$PA \vdash \forall u \notin B_n \neg Tr_n(u).$$

Further, let  $Prf_u(m, x)$  denote the primitive recursive formula expressing the predicate "  $m$  is (the gödel number of) a proof of the formula  $x$  from axioms of PA together with the additional axiom  $u$  ".



Given an arithmetic sentence  $B$   $\text{Prf}_B(m,x)$  will abbreviate the formula  $\text{Prf}_{\ulcorner B \urcorner}(m,x)$  and  $\text{Prf}(m,x)$  will denote  $\text{Prf}_{0=0}(m,x)$ .

By the routine formalization of the definition below with the aid of the Fixed Point Theorem one can define an arithmetic  $\Lambda_{n+1}$ -function  $h$  such that PA proves that  $h(0) = 0$  and for all  $m$   $h(m+1)$  is "computed" by the following instructions:

if  $z > h(m)$ ,  $z > 1$  and  $\text{Prf}(m+1, \ulcorner l \neq z \urcorner$ ) (1)

then  $h(m+1) = z$

else if  $h(m) = 0$  and  $\exists y_1, y_2 \leq m+1$  ( $\text{Prf}(y_1, \ulcorner l \neq 1 \urcorner$ )  $\wedge$   $\text{Prf}(y_2, \ulcorner l \neq 2 \urcorner$ ) ) (2)

then if  $\exists z, u \leq m+1$  ( $\text{Tr}_n(u) \wedge \text{Prf}_u(z, \ulcorner l \neq 1 \urcorner$ )  $\wedge$   $\forall t, w < z$  ( $\neg \text{Tr}_n(w) \vee \neg \text{Prf}_w(t, \ulcorner l \neq 2 \urcorner$ ))) (3)

then  $h(m+1) = 1$   
else  $h(m+1) = 2$  (4)

else  $h(m+1) = h(m)$ . (5)

Here  $l=z$  denotes the formula  $\exists N \forall m > N h(m) = z$ .

The following Lemma establishes the properties of our function  $h$  similar to those of the original Solovay's function (cf [8],[3]).

LEMMA 2.

1.  $\text{PA} \vdash \forall m ( h(m) \leq h(m+1) )$ ;  $\text{PA} \vdash l=0 \vee l=1 \vee \dots \vee l=k$ ;
2.  $i, j \in K$  and  $i \neq j \Rightarrow \text{PA} \vdash l=i \rightarrow l \neq j$ ;
3.  $j > i > 0 \Rightarrow \text{PA} \vdash l=i \rightarrow \neg \text{Pr} \ulcorner l \neq j \urcorner$ ;
4.  $i > 0 \Rightarrow \text{PA} \vdash l=i \rightarrow \text{Pr} \ulcorner l > i \urcorner$ ;
5.  $\text{PA} + l \in \{1, 2\}$  is a consistent theory;
6.  $l=0$  is true;
7.  $\text{PA} \vdash l \in \{1, 2\} \leftrightarrow \text{Pr} \ulcorner l \notin \{1, 2\} \urcorner \wedge \forall i > 1 \neg \text{Pr} \ulcorner l \neq i \urcorner$ .

PROOF. Claims 1 and 2 follow immediately from the definition of  $h$ . Statement 3 is proved as in Solovay's paper; the following argument can be formalized in PA: "If  $l \neq j$  is provable then it has PA-proofs with arbitrarily large gödel numbers. So, if  $l=i$  then  $h$  is bound to make a move from  $i$  by clause (1) because  $j > i > 0$ ".

The proof of 4 is not quite the same as in [8], because our function  $h$  is not  $\Sigma_1$ . Yet for all  $i > 1$  one can easily show that the sentence  $\exists m h(m) = i$  is  $\Sigma_1^{\text{PA}}$ .

In fact,

$$\text{PA} \vdash \exists m h(m)=i \leftrightarrow \exists m (\text{Prf}(m, \ulcorner l \neq i \urcorner) \wedge \forall y < m \forall j \in K (\neg j < i \wedge \text{Prf}(y, \ulcorner l \neq j \urcorner) \rightarrow \exists z < i \exists x < y (\text{Prf}(x, \ulcorner l \neq z \urcorner) \wedge \neg z < j) ) ).$$

The ( $\rightarrow$ ) implication follows from the monotonicity of  $h$ , for if  $m$  is the least (gödel number of a) proof of  $l \neq i$ ,  $h(y)$  is to be beneath  $i$  for all  $y < m$ . So, by clause (1), if at some stage  $y < m$  we receive a proof of  $l \neq j$ , where not  $j < i$ ,  $h(y)$  is already to be not  $< j$ . And this is possible only if we have earlier received a proof of some sentence  $l \neq z$ , where  $z < i$  but not  $z < j$ .

In order to prove the converse implication we only have to check that  $h$  cannot make a move from beneath  $i$  before stage  $m$ . Since both 1 and 2 are beneath  $i$ , such a move can only be made by clause (1); but this is impossible because if  $h$  moved from some  $w < i$  to  $j$ , where not  $j < i$ , we would earlier have obtained a proof of  $l \neq z$  for some  $z < i$  and  $z$  not  $< j$ . By property (iv) of the model  $\mathcal{K}$ , either  $w < z$  or  $z < w$ . The latter is not the case since  $w < j$  but  $z$  is not. So,  $h$  is bound to have moved from  $w$  to  $z$  earlier than to  $j$ . Absurd.

Suppose now that  $i > 1$ . Then

$$\begin{aligned} \text{PA} \vdash l=i &\rightarrow \exists m h(m)=i \\ &\rightarrow \text{Pr} \ulcorner \exists m h(m)=i \urcorner \\ &\rightarrow \text{Pr} \ulcorner l > i \urcorner. \end{aligned}$$

But for all  $i > 0$  trivially

$$\text{PA} \vdash l=i \rightarrow \text{Pr} \ulcorner l \neq i \urcorner;$$

hence

$$\text{PA} \vdash l=i \rightarrow \text{Pr} \ulcorner l > i \urcorner.$$

We treat the case  $i \in \{1, 2\}$  in a different way. By clause (2),  $h$  is allowed to make a move to 1 only after having received both proofs of  $l \neq 1$  and  $l \neq 2$ . It follows that

$$\text{PA} \vdash l=1 \rightarrow \text{Pr} \ulcorner l \notin \{1, 2\} \urcorner.$$

We also have

$$\text{PA} \vdash \text{Pr} \ulcorner l \notin \{1, 2\} \urcorner \rightarrow l \neq 0;$$

hence

$$\begin{aligned} \text{PA} \vdash l=1 &\rightarrow \text{Pr} \ulcorner l \notin \{1, 2\} \urcorner \\ &\rightarrow \text{Pr} \ulcorner \text{Pr} \ulcorner l \notin \{1, 2\} \urcorner \urcorner \\ &\rightarrow \text{Pr} \ulcorner l \neq 0 \urcorner. \end{aligned}$$

Thus we obtain

$$\begin{aligned} \text{PA} \vdash l=1 &\rightarrow \text{Pr} \ulcorner l \notin \{0, 1, 2\} \urcorner \\ &\rightarrow \text{Pr} \ulcorner l > 1 \urcorner. \end{aligned}$$

The case  $i=2$  is fully symmetrical; so statement 4 follows.

Claim 6 is trivial.

To check 5 and 7 notice that if PA proves  $l \notin \{1,2\}$ , then  $h$  is bound to make a move from 0.

REMARK. Note that statement 3 of this Lemma is weaker than the corresponding statement of Solovay, because we have not proved that

$$\text{PA} \vdash l=0 \rightarrow \neg \text{Pr} \ulcorner l \neq 1 \urcorner.$$

Although it might seem that our function  $h$  lacked this property, because  $h$  was to stay in 0 until we received *both* proofs of  $l \neq 1$  and  $l \neq 2$ , nevertheless one can show within PA that once we have received one of these proofs we are guaranteed to receive the other. We shall put off the proof of this statement for it is needed only to improve slightly on a simpler result that we are going to obtain first.

Now we define an arithmetical interpretation  $f$  exactly as in Solovay's paper:

$$f(p) := \exists i (l=i \wedge i \vDash p).$$

As in [8] we obtain the following Corollary.

COROLLARY 2. For all nodes  $i > 0$  and all formulae  $\psi$

$$i \vDash \psi \Rightarrow \text{PA} \vdash l=i \rightarrow f(\psi).$$

LEMMA 3. Suppose  $A \in \mathbf{B}_n$  is an arithmetic sentence, such that the theory  $\text{PA} + l \in \{1,2\} + A$  is consistent. Then  $\text{PA} + l \in \{1,2\} + A$  does not prove either  $l=1$  or  $l \neq 1$ .

PROOF. Since  $\text{PA} + l \in \{1,2\} \vdash l=1 \leftrightarrow l \neq 2$  and 1 and 2 are symmetrical, we only have to prove that  $\text{PA} + l \in \{1,2\} + A$  does not prove  $l \neq 1$ .

Suppose for a reductio that it does. Let  $m$  be the least ( gödel number of a ) proof of  $l \neq 1$  from an arithmetic sentence  $B \in \mathbf{B}_n$  consistent with  $\text{PA} + l \in \{1,2\}$ , i.e.

$\text{Prf}_B(m, \ulcorner l \neq 1 \urcorner)$  ,  $\text{PA} + l \in \{1,2\} + B$  is a consistent theory

and for all  $y < m$  and all  $C \in \mathbf{B}_n$

if  $\text{Prf}_C(y, \ulcorner l \neq 1 \urcorner)$  then  $\text{PA} + l \in \{1,2\} + C$  is inconsistent.

Such a  $B$  exists because by Lemma 2.7  $l \in \{1,2\}$  is  $\mathbf{B}_1^{\text{PA}}$ , hence  $A \wedge l \in \{1,2\}$  is  $\mathbf{B}_n^{\text{PA}}$ . Further we obtain

$$\text{PA} + l \in \{1,2\} + B \vdash \text{Tr} \ulcorner B \urcorner \wedge \text{Prf}_B(m, \ulcorner l \neq 1 \urcorner)$$

$$\vdash l=1 \vee \exists y < m \exists u < m (\text{Tr}_n(u) \wedge \text{Prf}_u(y, \ulcorner l \neq 2 \urcorner)).$$

Reason in  $PA + \ell \in \{1,2\} + B$  : "Let  $t$  be the least stage such that we have received both proofs of  $\ell \neq 1$  and  $\ell \neq 2$  by the moment  $t$ . So, since  $\ell \in \{1,2\}$  the function  $h$  is not allowed to leave 0 before  $t$ . Thus, by clause (2),  $h$  is to make a move either to 1 or to 2. If not  $\exists y < m \exists u < m (Tr_n(u) \wedge Prf_u(y, \lceil \ell \neq 2 \rceil))$  then  $h$  will move to 1 by clause (3) because  $B$  is true  $B_n$ ."

Further

$PA + \ell \in \{1,2\} + B \vdash \exists y < m \exists u < m (Tr_n(u) \wedge Prf_u(y, \lceil \ell \neq 2 \rceil))$ ,  
because by our assumption

$$PA + B \vdash \ell \neq 1.$$

We claim that there is a sentence  $C \in B_n$  such that  $C$  is consistent with  $PA + \ell \in \{1,2\}$  and for some  $y < m$   $Prf_C(y, \lceil \ell \neq 2 \rceil)$ .

Otherwise we would have for all  $y, u < m$  that

either not  $Prf_u(y, \lceil \ell \neq 2 \rceil)$  and then  $PA \vdash \neg Prf_u(y, \lceil \ell \neq 2 \rceil)$ ;

or  $u$  is not a gödel number of a  $B_n$ -formula and then

$$PA \vdash \neg Tr_n(u);$$

or  $u = \lceil C \rceil$ ,  $Prf_u(y, \lceil \ell \neq 2 \rceil)$ ,  $C \in B_n$  and  $PA + \ell \in \{1,2\} \vdash \neg C$ ; then

$$PA + \ell \in \{1,2\} \vdash \neg Tr_n \lceil C \rceil.$$

In each case we obtain

$$PA + \ell \in \{1,2\} \vdash \neg Tr_n(u) \vee \neg Prf_u(y, \lceil \ell \neq 2 \rceil).$$

Thus,

$$PA + \ell \in \{1,2\} \vdash \forall y < m \forall u < m (\neg Tr_n(u) \vee \neg Prf_u(y, \lceil \ell \neq 2 \rceil)).$$

It follows that  $PA + \ell \in \{1,2\} + B$  is inconsistent. Absurd.

Applying now similar arguments to  $C$  and  $y$  instead of  $B$  and  $m$  respectively we obtain a sentence  $D \in B_n$  such that  $D$  is consistent with  $PA + \ell \in \{1,2\}$  and for some  $x < y < m$   $Prf_D(x, \lceil \ell \neq 1 \rceil)$ . This contradicts our assumption that  $m$  is the minimal gödel number of a proof of  $\ell \neq 1$  from a  $B_n$ -sentence consistent with  $PA + \ell \in \{1,2\}$ .

COROLLARY 3. For no sentence  $A \in B_n$

$$PA + \ell \in \{1,2\} \vdash A \leftrightarrow \ell = 1.$$

PROOF. Suppose that

$$PA + \ell \in \{1,2\} \vdash A \leftrightarrow \ell = 1.$$

Then by Lemma 3 the theory  $PA + \ell \in \{1,2\} + A$  is inconsistent.

It follows that

$$\text{PA} + \ell \in \{1, 2\} \vdash \ell \neq 1.$$

Applying Lemma 3 again we obtain that  $\text{PA} + \ell \in \{1, 2\}$  is inconsistent. This contradicts Lemma 2.5.

Now we are in a position to prove the following weaker version of our theorem.

PROPOSITION. Suppose  $\phi$  is a formula of  $\mathcal{L}$  such that for no boolean combination  $\psi$  of boxed formulae  $\text{GL} \vdash \phi \leftrightarrow \psi$ . Then for each  $n \geq 1$  there exists an arithmetical interpretation  $f$  such that  $f(\phi) \notin \mathbf{B}_n^{\text{PA}}$ .

PROOF. Let the model  $\mathcal{K}$  and the function  $h$  be defined as above. By Corollary 2

$$\text{PA} \vdash \ell=1 \rightarrow f(\phi) \quad \text{and} \quad \text{PA} \vdash \ell=2 \rightarrow \neg f(\phi).$$

So,

$$\text{PA} + \ell \in \{1, 2\} \vdash \ell=1 \leftrightarrow f(\phi).$$

If for some  $A \in \mathbf{B}_n$

$$\text{PA} + \ell \in \{1, 2\} \vdash A \leftrightarrow f(\phi),$$

we shall get

$$\text{PA} + \ell \in \{1, 2\} \vdash \ell=1 \leftrightarrow A$$

contradicting Corollary 3.

Of course, this Proposition is weaker than the Theorem because it does not give us any upper bound on the complexity of the interpretation constructed.

To improve the result in this respect we have to do an additional piece of work.

LEMMA 4. Suppose  $\phi$  is a nonstable formula. Then there is a model  $\mathcal{K} = (\mathcal{K}, <, \Vdash)$  such that

- (i)  $\mathcal{K} = \{0, \dots, k\}$  for some  $k \in \mathbb{N}$ ;
- (ii) 0 is the bottom node of  $\mathcal{K}$ ;
- (iii) 1 and 2 have the same successors and predecessors;
- (iv) for all  $i, z, w \in \mathcal{K}$ , if  $1 < z, w$  and  $z, w < i$  then either  $z < w$  or  $w < z$ ;
- (v) the set of predecessors of 1 and 2 is linearly ordered;
- (vi) for all subformulae  $\psi$  of formulae  $\phi$   $0 \Vdash \Box \psi \rightarrow \psi$ ;
- (vii)  $1 \Vdash \phi$  and  $2 \Vdash \neg \phi$ .

PROOF. Let  $\mathcal{K}_1$  and  $\mathcal{K}_2$  be the given pair of (treelike) models such that  $\mathcal{K}_1 \Vdash \phi$  and  $\mathcal{K}_2 \Vdash \neg \phi$ . We identify the corresponding nodes of  $\mathcal{K}_1$  and  $\mathcal{K}_2$  except the roots and add a sufficiently long finite

linearly ordered set  $T$  below them. The words "sufficiently long" mean that  $T$  should be longer than the number of all boxed subformulae of  $\phi$ . The forcing relation of propositional variables is defined arbitrarily on  $T$ .

By the Pigeon-Hole Principle there is a node  $x \in T$  such that the same boxed subformulae of  $\phi$  are forced at  $x$  as at the immediate predecessor of  $x$ . Clearly we shall have

$$x \Vdash \Box\psi \rightarrow \psi$$

for all subformulae  $\Box\psi$  of formula  $\phi$ . Thus we may take  $x$  as the bottom node of the model required.

Let  $\mathcal{K} = (K, <, \Vdash)$  be a model as in Lemma 4. We modify the definition of our function  $h$  as follows:  $h(0)=0$  and for all  $m$

$$\text{if } z > h(m), z \notin \{1,2\} \text{ and } \text{Prf}(m+1, \ulcorner l \neq z \urcorner) \quad (1)$$

$$\text{then } h(m+1)=z$$

$$\text{else if } h(m) < 1 \text{ and } \exists y_1, y_2 \leq m+1 ( \text{Prf}(y_1, \ulcorner l \neq 1 \urcorner) \wedge \text{Prf}(y_2, \ulcorner l \neq 2 \urcorner) ) \quad (2)$$

$$\text{then if } \exists z, u \leq m+1 ( \text{Tr}_n(u) \wedge \text{Prf}_u(z, \ulcorner l \neq 1 \urcorner) \wedge \forall t, w < z ( \neg \text{Tr}_n(w) \vee \neg \text{Prf}_w(t, \ulcorner l \neq 2 \urcorner) ) ) \quad (3)$$

$$\text{then } h(m+1)=1$$

$$\text{else } h(m+1)=2 \quad (4)$$

$$\text{else } h(m+1)=h(m). \quad (5)$$

It is easy to see that all (the proofs of) statements of Lemma 2 except Statement 4 remain true for the modified function  $h$ . Now we are going to show that even a stronger version of 4 holds. The following technical claim is a generalized variant of Lemma 4.1.6 of [7] ( cf also Corollary 2.23 [9] ).

LEMMA 5. For all arithmetic sentences  $A \in \mathbf{B}_n$

$$\text{PA} \vdash \forall m \text{Pr} \ulcorner \exists x < m \exists u < m ( \text{Prf}_u(x, \ulcorner A \urcorner) \wedge \text{Tr}_n(u) ) \rightarrow A \urcorner$$

PROOF. Since  $\text{Pr}$  commutes with restricted quantifiers it is sufficient to prove that

$$\text{PA} \vdash \forall m \forall u < m \text{Pr} \ulcorner \exists x < m ( \text{Prf}_u(x, \ulcorner A \urcorner) \wedge \text{Tr}_n(u) ) \rightarrow A \urcorner$$

Clearly

$$\begin{aligned} \text{PA} \vdash \exists x < m \text{Prf}_u(x, \ulcorner A \urcorner) \wedge u \in \mathbf{B}_n &\rightarrow ( \text{Pr}(u \dot{\rightarrow} \ulcorner A \urcorner) \wedge u \in \mathbf{B}_n ) \\ &\rightarrow ( \text{Pr}(u \dot{\rightarrow} \ulcorner A \urcorner) \wedge \text{Pr}(\ulcorner \text{Tr}_n(u) \urcorner \dot{\leftarrow} u) ) \\ &\rightarrow \text{Pr} \ulcorner \text{Tr}_n(u) \rightarrow A \urcorner \end{aligned}$$

$$\rightarrow \text{Pr} \lceil \exists x < m (\text{Prf}_u(x, \lceil A \rceil) \wedge \text{Tr}_n(u)) \rightarrow A \rceil.$$

On the other hand, since  $\text{PA} \vdash u \notin \mathbf{B}_n \rightarrow \neg \text{Tr}_n(u)$

$$\begin{aligned} \text{PA} \vdash \forall x < m \neg \text{Prf}_u(x, \lceil A \rceil) \vee u \notin \mathbf{B}_n &\rightarrow \text{Pr} \lceil \forall x < m \neg \text{Prf}_u(x, \lceil A \rceil) \vee u \notin \mathbf{B}_n \rceil \\ &\rightarrow \text{Pr} \lceil \exists x < m (\text{Prf}_u(x, \lceil A \rceil) \wedge \text{Tr}_n(u)) \rightarrow \perp \rceil \\ &\rightarrow \text{Pr} \lceil \exists x < m (\text{Prf}_u(x, \lceil A \rceil) \wedge \text{Tr}_n(u)) \rightarrow A \rceil. \end{aligned}$$

Thus  $\text{PA} \vdash \forall m \forall u \text{Pr} \lceil \exists x < m (\text{Prf}_u(x, \lceil A \rceil) \wedge \text{Tr}_n(u)) \rightarrow A \rceil$  and Lemma 5 follows.

LEMMA 6.  $\text{PA} \vdash \text{Pr} \lceil l \neq 1 \rceil \leftrightarrow \text{Pr} \lceil l \neq 2 \rceil$ .

PROOF. It is easy to see that

$$\text{PA} \vdash l > 1 \rightarrow \exists i > 1 \text{Pr} \lceil l \neq i \rceil.$$

Besides,

$$\text{PA} \vdash l < 1 \wedge \text{Pr} \lceil l \neq 1 \rceil \rightarrow \neg \text{Pr} \lceil l \neq 2 \rceil,$$

because clearly  $h$  would make a move from beneath 1, if we received both proofs of  $l \neq 1$  and  $l \neq 2$ . Hence

$$\begin{aligned} \text{PA} \vdash \text{Prf}(m, \lceil l \neq 1 \rceil) &\rightarrow \text{Pr} \lceil \text{Pr} \lceil l \neq 1 \rceil \rceil \\ &\rightarrow \text{Pr} \lceil l < 1 \rightarrow \neg \text{Pr} \lceil l \neq 2 \rceil \rceil. \end{aligned}$$

On the other hand, by clause (3) of the definition of  $h$   
 $\text{PA} \vdash \text{Prf}(m, \lceil l \neq 1 \rceil) \rightarrow (l = 2 \rightarrow \exists x < m \exists u < m (\text{Prf}_u(x, \lceil l \neq 2 \rceil) \wedge \text{Tr}_n(u)))$ ,  
hence

$$\text{PA} \vdash \text{Prf}(m, \lceil l \neq 1 \rceil) \rightarrow \text{Pr} \lceil l = 2 \rightarrow \exists x < m \exists u < m (\text{Prf}_u(x, \lceil l \neq 2 \rceil) \wedge \text{Tr}_n(u)) \rceil.$$

Thus,

$$\begin{aligned} \text{PA} \vdash \text{Prf}(m, \lceil l \neq 1 \rceil) &\rightarrow \text{Pr} \lceil l > 1 \vee l < 1 \vee l = 2 \rceil \\ \rightarrow \text{Pr} \lceil (\exists i > 1 \text{Pr} \lceil l \neq i \rceil) \vee \neg \text{Pr} \lceil l \neq 2 \rceil \vee \exists x < m \exists u < m (\text{Prf}_u(x, \lceil l \neq 2 \rceil) \wedge \text{Tr}_n(u)) \rceil. \end{aligned}$$

Obviously

$$\text{PA} \vdash \text{Pr} \lceil (\exists i > 1 \text{Pr} \lceil l \neq i \rceil) \rightarrow l \neq 2 \rceil$$

and by Lemma 5

$$\text{PA} \vdash \forall m \text{Pr} \lceil \exists x < m \exists u < m (\text{Prf}_u(x, \lceil l \neq 2 \rceil) \wedge \text{Tr}_n(u)) \rightarrow l \neq 2 \rceil.$$

It follows that

$$\begin{aligned} \text{PA} \vdash \text{Prf}(m, \lceil l \neq 1 \rceil) &\rightarrow \\ \rightarrow \text{Pr} \lceil \text{Pr} \lceil l \neq 2 \rceil \rightarrow (\exists i > 1 \text{Pr} \lceil l \neq i \rceil) \vee \exists x < m \exists u < m (\text{Prf}_u(x, \lceil l \neq 2 \rceil) \wedge \text{Tr}_n(u)) \rceil \\ \rightarrow \text{Pr} \lceil \text{Pr} \lceil l \neq 2 \rceil \rightarrow l \neq 2 \rceil. \end{aligned}$$

By formalized Löb's theorem we obtain

$$PA \vdash \text{Prf}(m, \lceil \ell \neq 1 \rceil) \rightarrow \text{Pr} \lceil \ell \neq 2 \rceil,$$

whence

$$PA \vdash \exists m \text{Prf}(m, \lceil \ell \neq 1 \rceil) \rightarrow \text{Pr} \lceil \ell \neq 2 \rceil.$$

The converse implication is symmetrical, so Lemma 6 follows.

COROLLARY 5. For all  $i, j \in K$

$$i < j \Rightarrow PA \vdash \ell = i \rightarrow \neg \text{Pr} \lceil \ell \neq j \rceil.$$

PROOF. This is trivial for  $j \notin \{1, 2\}$ . Suppose  $j=1$ . The following argument can be formalized in PA:

"If  $\text{Pr} \lceil \ell \neq 1 \rceil$  then by Lemma 6  $\text{Pr} \lceil \ell \neq 2 \rceil$ , hence we are to receive both proofs of  $\ell \neq 1$  and  $\ell \neq 2$  by some stage  $m$ . Since  $\ell < 1$  condition (2) will be satisfied, so  $h$  is bound to make a move either to 1 or to 2. Absurd."

We define the arithmetical interpretation  $f$  just as before and as in the proof of Solovay's second theorem (cf [8]) obtain the following

COROLLARY 6. For all subformulae  $\psi$  of formula  $\phi$

$$PA \vdash f(\psi) \leftrightarrow \exists i (\ell = i \wedge i \vDash \psi).$$

Consequently  $f(\phi)$  is equivalent to a boolean combination of formulae of the form  $\ell = i$ , because the quantifier  $\exists i$  actually ranges over the finite set  $K$ . For  $i \notin \{1, 2\}$  it is easily seen that the formula  $\ell = i$  is  $B_1^{\text{PA}}$ . Now for  $i \in \{1, 2\}$ .

It follows immediately from the definition of  $h$  that

$$PA \vdash \ell = 1 \leftrightarrow \ell \in \{1, 2\} \wedge \exists m (\exists u < m (\text{Tr}_n(u) \wedge \text{Prf}_u(m, \lceil \ell \neq 1 \rceil) \wedge \forall y < m \forall w < m (\neg \text{Tr}_n(w) \vee \neg \text{Prf}_w(y, \lceil \ell \neq 2 \rceil) ))).$$

We have already seen that  $\ell \in \{1, 2\}$  is  $B_1^{\text{PA}}$ , hence  $\ell = 1$  is  $\Sigma_{n+1}^{\text{PA}}$ . But clearly

$$PA \vdash \ell = 1 \leftrightarrow \ell \neq 2 \wedge \ell \in \{1, 2\},$$

whence by symmetry of nodes 1 and 2  $\ell = 1$  is also  $\Pi_{n+1}^{\text{PA}}$ . Thus  $\ell = 1$  and  $f(\phi)$  both are  $\Delta_{n+1}^{\text{PA}}$ .

We have received the desired upper bound on the arithmetical complexity of  $f(\phi)$ . To obtain the lower one we proceed just as in the proof of Proposition, because evidently Lemma 3 still holds for the modified function  $h$ .

I would like to thank D.de Jongh for useful discussions and V.Yu. Shavrukov for a number of comments on a draft of this paper.



## REFERENCES

1. C.Kent. The relation of  $A$  to  $\text{Prov}^{\lceil A \rceil}$  in the Lindenbaum sentence algebra. *Journal of Symbolic Logic*, 1973, v.38, pp.295-298.
2. D.Guaspari. Sentences implying their own provability. *Journal of Symbolic Logic*, 1983, v.48, pp.777-789.
3. C.Smoryński. *Self-Reference and Modal Logic*, Springer, 1985.
4. A.Visser. *A Course in Bimodal Provability Logic*. Logic Group Preprint Series No.20, 1987, Department of Philosophy, University of Utrecht, Heidelberglaan 2, 3584CS Utrecht.
5. D.de Jongh, D.Pianigiani. Solution of a problem of David Guaspari, manuscript, 1989, Amsterdam.
6. G.Kreisel, A.Lévy. Reflection principles and their use for establishing the complexity of axiom systems. *Z.math.Log.Grundl. Math.*, 1968, Bd.14, S.97-142.
7. C.Smoryński. Incompleteness theorems. In: *Handbook of Mathematical Logic*, J. Barwise (ed.), North-Holland, 1977.
8. R.Solovay. Provability interpretations of modal logic. *Israel J. Math.*, 1976, v.25, pp.287-304.
9. C.Bennet. On some orderings of extensions of arithmetic. Ph.D. Thesis, Department of Philosophy, University of Göteborg, Göteborg, 1986.



# 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

*Mathematical Logic and Foundations*

ML-90-01 Harold Schellinx

ML-90-02 Jaap van Oosten

*Computation and Complexity Theory*

CT-90-01 John Tromp, Peter van Emde Boas

*Other Prepublications*

X-90-01 A.S. Troelstra

X-90-02 Maarten de Rijke

X-90-03 L.D. Beklemishev

A Generalized Quantifier Logic for Naked Infinitives  
Dynamic Montague Grammar

Isomorphisms and Non-Isomorphisms of Graph Models  
A Semantical Proof of De Jongh's Theorem

Associative Storage Modification Machines

Remarks on Intuitionism and the Philosophy of Mathematics,  
Revised Version

Some Chapters on Interpretability Logic

On the Complexity of Arithmetical Interpretations of Modal Formulae

