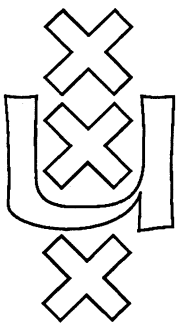


**Institute for Language, Logic and Information**

**THE LINDENBAUM FIXED POINT ALGEBRA  
IS UNDECIDABLE**

V.Yu. Shavrukov

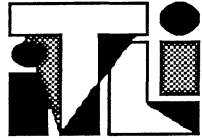
ITLI Prepublication Series  
X-90-07



**University of Amsterdam**

## The ITLI Prepublication Series

- 1986**
- 86-01 The Institute of Language, Logic and Information  
A Semantical Model for Integration and Modularization of Rules
- 86-02 Peter van Emde Boas Categorical Grammar and Lambda Calculus
- 86-03 Johan van Benthem A Relational Formulation of the Theory of Types
- 86-04 Reinhard Muskens Some Complete Logics for Branched Time, Part I Well-founded Time, Logical Syntax Forward looking Operators
- 86-05 Kenneth A. Bowen, Dick de Jongh
- 86-06 Johan van Benthem
- 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 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 Dynamic Data Structures
- Leen Torenvliet, Peter van Emde Boas
- 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  $\text{IA}_0 + \Omega_1$
- 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 Nondeterministic Complexity Classes
- Leen Torenvliet
- 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 Muidergracht 24  
1018TV Amsterdam

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

# **THE LINDENBAUM FIXED POINT ALGEBRA IS UNDECIDABLE**

V.Yu. Shavrukov  
Steklov Mathematical Institute  
Vavilova 42, 117966 Moscow GSP-1, U.S.S.R.



## The Lindenbaum Fixed Point Algebra Is Undecidable

V. Yu. Shavrukov

Abstract. We prove that the first order theory of the fixed point algebra corresponding to an r.e. consistent theory containing arithmetic is hereditarily undecidable.

Fixed point algebras (f.p.a.'s) were introduced by Smoryński in [4] (see also [6]). A f.p.a. is a pair (A,B) of Boolean algebras, where elements of B are mappings  $A \rightarrow A$ , satisfying the following conditions:

- (i)  $\alpha = \beta \iff \forall a \alpha a = \beta a$
- (ii)  $(\alpha \# \beta)a = \alpha a \# \beta a$  (# is a Boolean operation)
- (iii)  $\forall a \exists \alpha \forall b \alpha b = a$
- (iv)  $\forall \alpha \exists a \alpha a = a$

(elements of A are denoted by Latin letters and those of B by Greek ones).

F.p.a.'s arise most naturally from the consideration of a theory T containing arithmetic. Take A to be the Lindenbaum sentence algebra of T (i.e. the set of sentences of T modulo T-provable equivalence). Call a formula F(x)

with no variable other than  $x$  free extensional if  $\mathbb{T} \vdash \varphi \leftrightarrow \psi$  implies  $\mathbb{T} \vdash F(\overline{\varphi}) \leftrightarrow F(\overline{\psi})$  for all sentences  $\varphi$  and  $\psi$ . These formulas induce mappings  $\tilde{F}: \varphi \mapsto F(\overline{\varphi})$  on  $A$ . Let  $B$  be the collection of mappings of this form. Property (iv) is then an algebraic translation of Gödel's Diagonal Lemma. These algebras were studied by Montagna [3] and Solovay [7]. In particular the latter paper establishes the (recursive) isomorphism of f.p.a.'s associated with r.e. consistent theories containing Peano Arithmetic (PA). We shall call this f.p.a. the Lindenbaum f.p.a.

Theorem. The first order theory of the Lindenbaum f.p.a. is hereditarily undecidable.

(By "hereditarily" we mean that every one of its subtheories is also undecidable).

A first order theory  $S$  possessing finite models is said to be finitely inseparable if the set of theorems of  $S$  is effectively inseparable from the set of sentences false in an appropriate finite model of  $S$ . Finite inseparability of a finitely axiomatized theory  $S$  clearly implies that the theory of finite models of  $S$  is hereditarily undecidable. The finite inseparability of the theory of f.p.a.'s is shown in [5].

We also recall a theorem of Taiclin [8] which asserts that the theory of partially ordered sets is finitely inseparable (see also [2]). Consider the following interpretation  $+$  of the language of poset theory into the language of f.p.a.'s: The domain of  $+$  is the set of fixed points of  $\alpha$ , where  $\alpha$  is a parameter from  $B$ , and the

ordering relation  $x \leq y$  is interpreted by  $\mathbb{T}$ -provable implication. The lemma below says that every finite poset is isomorphic to the structure induced by the Lindenbaum f.p.a. via this interpretation provided that we choose a suitable  $\alpha$ . Then the set of (poset) sentences

$$\{ A \mid \forall \alpha A^+ \text{ holds in the Lindenbaum f.p.a.} \}$$

constitutes a subtheory of the theory of finite posets and is therefore hereditarily undecidable. The theorem follows. The above construction falls within the general scheme of the relative elementary definability method detailed in [1].

Lemma. Let  $R$  be a (reflexive) partial order on  $n > 0$ . Then there exists an  $(\Sigma_1^0)$  extensional formula  $B(x)$  such that  $(n, R)$  is isomorphic to  $(\{ a \in A \mid \tilde{B}a = a \}, \text{provable implication})$ .

Proof. By Solovay's result we may assume that we are working with PA.

We shall define within this theory a recursive procedure which in the course of its operation may paint (gödelnumbers of) some sentences black. The arithmetic formula  $B(x)$  will say "x is eventually painted black". The procedure is defined by stages using the gödelnumber of  $B(x)$ ; this is justified by the formalized recursion theorem. For technical convenience we assume that the language of arithmetic contains two distinct biconditionals  $\leftrightarrow$  and  $\equiv$ . We also fix an enumeration of theorems of PA in which each theorem occurs infinitely often.

Stage 0. Pick  $n$  fixed points  $\phi_0, \dots, \phi_{n-1}$  of  $B(x)$  with distinct gödelnumbers. Now take  $n$  empty boxes and place  $\phi_i$  in the  $i$ th box. Go to the next stage.

Stage  $m+1$ . For every arithmetic sentence  $\psi$  let  $Y_\psi$  denote the finite set

$$\{\psi\} \cup \{ \varphi \mid \bar{\varphi} \leq m \text{ and } \varphi \leftrightarrow \psi \text{ is a tautological consequence of the first } m \text{ theorems of PA} \}.$$

Look at the  $m$ th theorem of PA.

Case 1. The  $m$ th theorem reads:  $\psi \leftrightarrow \phi_i$ .

Put every element of  $Y_\psi$  that has not been placed in some box at an earlier stage into the  $i$ th box. Go to the next stage.

Case 2. The  $m$ th theorem reads:  $\psi \equiv B(\bar{\psi})$  and  $\psi$  is not yet in any box.

Look for a sentence in  $Y_\psi$  which has already been put in some box. Let  $\phi$  in the  $j$ th box be the first such found. Then put all elements of  $Y_\psi$  into the  $j$ th box except those put into some box at preceding stages. In case no such  $\phi$  exists put all of  $Y_\psi$  into any box you like. Go to the next stage.

Case 3. The  $m$ th theorem reads:  $\phi_i \rightarrow \phi_j$  and  $i \text{ nonR } j$ .

Paint the (contents of the)  $k$ th box black whenever  $iRk$ . Then halt the whole procedure.

Case 4. None of the previous cases apply. Go to the next stage.



The definition of the procedure is now complete. Note that the instructions of Case 2 are recursive because at every stage each box contains only finitely many sentences. We proceed to prove that  $B(x)$  is as required. The proof is contained in the following sequence of claims.

a) (PA) At any stage any sentence can be found in at most one box.

Obvious by induction on stages.

b) Case 3 never happens.

Suppose it does:  $i \text{ nonR } j$  and the  $m$ th theorem of PA is  $\phi_i \rightarrow \phi_j$ . Reason in PA: Since  $iRi$ , by the instructions of Case 3  $\phi_i$  is painted black whereas  $\phi_j$  is in the  $j$ th box and therefore by (a) not in any  $k$ th box with  $iRk$ . Hence  $\phi_j$  is never painted black for there are no stages after Case 3. We have proved  $B(\overline{\phi_i})$  and not  $B(\overline{\phi_j})$ . On the other hand,  $\phi$ 's are fixed points of  $B(x)$  and  $\phi_i \rightarrow \phi_j$ , ergo  $B(\overline{\phi_i}) \rightarrow B(\overline{\phi_j})$  which is a contradiction. Now recall that PA is consistent.

c) For every  $m \in \omega$  there is a Stage  $m+1$ .

Immediate from (b).

d) PA  $\vdash \phi_i \rightarrow \phi_j$  implies  $iRj$ .

Suppose  $i \text{ nonR } j$  and  $\phi_i \rightarrow \phi_j$  is the  $m$ th theorem of PA. At Stage  $m+1$  which exists by (c) we would have Case 3. But this contradicts (b).

e)  $iRj$  implies PA  $\vdash \phi_i \rightarrow \phi_j$ .

Reason in PA: The only way something can be painted black is via Case 3. Suppose the (contents of the)  $i$ th box is

painted black. But since  $iRj$  and  $R$  is transitive so is the  $j$ th box. Therefore  $B(\overline{\phi_i}) \rightarrow B(\overline{\phi_j})$ . Infer  $\phi_i \rightarrow \phi_j$  by the choice of  $\phi$ 's. In case nothing is ever painted black we clearly have not  $\phi_i$  and not  $\phi_j$  which also implies  $\phi_i \rightarrow \phi_j$ .

f) For every sentence  $\varphi$  in the  $i$ th box

$$PA \vdash B(\overline{\varphi}) \leftrightarrow B(\overline{\phi_i}).$$

Let  $\varphi$  get into the  $i$ th box at Stage  $m$ . Since Case 3 has not yet happened it is verifiable in PA that  $\varphi$  is painted black iff  $\phi_i$  ever is.

g) For every  $i < n$  and every sentence  $\varphi$  in the  $i$ th box

$$PA \vdash \varphi \leftrightarrow B(\overline{\varphi}).$$

Suppose  $\varphi$  gets into the box because of Case I. We then have:

$$\begin{aligned} PA \vdash \varphi &\leftrightarrow \phi_i \\ &\leftrightarrow B(\overline{\phi_i}) && \text{by the choice of } \phi \text{'s} \\ &\leftrightarrow B(\overline{\varphi}) && \text{by (f)}. \end{aligned}$$

In case  $\varphi$  is in the  $i$ th box via Case 2 there is a  $\psi$  not yet in any box with  $PA \vdash \varphi \leftrightarrow \psi$  and  $PA \vdash \psi \equiv B(\overline{\psi})$ . The instructions of Case 2 insure that  $\psi$  is also placed in the  $i$ th box. Therefore

$$\begin{aligned} PA \vdash \varphi &\leftrightarrow \psi \\ &\leftrightarrow B(\overline{\psi}) \\ &\leftrightarrow B(\overline{\phi_i}) && \text{by (f)} \\ &\leftrightarrow B(\overline{\varphi}) && \text{by (f)}. \end{aligned}$$

h) Any two sentences in the same box are provably equivalent.

Immediate from (f) and (g).

i) Sentences from distinct boxes are never provably equivalent.

Denying this, we get  $PA \vdash \phi_i \leftrightarrow \phi_j$  for  $i \neq j$  by (h).

Applying (d) one has  $iRjRi$ . Quod non.

j) If  $PA \vdash \psi \leftrightarrow \phi_i$  then  $\psi$  is eventually put into the  $i$ th box.

Let  $\psi \leftrightarrow \phi_i$  be the  $m$ th theorem with  $m \geq \bar{\psi}$ . By (i)  $\psi$  can be in no box but the  $i$ th one. The instructions of Case 1 put  $\psi$  there at Stage  $m+1$  if not earlier.

k)  $PA \vdash \varphi \leftrightarrow \psi$  implies  $PA \vdash B(\bar{\varphi}) \leftrightarrow B(\bar{\psi})$ .

Consider first the case in which  $PA \vdash \varphi \leftrightarrow \phi_i$  for some  $i$ . By (j) both  $\varphi$  and  $\psi$  will eventually be put into the  $i$ th box. Now apply (f). Let  $PA \vdash \varphi \leftrightarrow \phi_i$  for no  $i$ . By (h) neither  $\varphi$  nor  $\psi$  is ever in any box. Let  $\varphi \leftrightarrow \psi$  be the  $m$ th theorem with  $m \geq \bar{\varphi}$ ,  $m \geq \bar{\psi}$ . At Stage  $m+1$  these sentences are not yet in a box and at succeeding stages for every sentence  $\theta$  we have  $\varphi \in Y_\theta$  iff  $\psi \in Y_\theta$ . Therefore it can be seen from inside PA that  $\varphi$  and  $\psi$  can only get into the same box and at the same stage. This means that their prospects of getting painted are precisely identical. The claim follows.

l) Every fixed point of  $B(x)$  is eventually placed in some box.

See Case 2.

m) Every fixed point of  $B(x)$  is provably equivalent to  $\phi_i$  for some  $i$ .

This follows from (l) and (h).

Note that (k) asserts the extensionality of  $B(x)$  while (d), (e) and (m) establish the desired isomorphism. The proof of the lemma and hence that of the theorem are now complete.

References

- [ 1 ] Ю.Л.Ершов, Проблемы разрешимости и конструктивные модели, Наука, Москва, 1980.
- [ 2 ] Ю.Л.Ершов, И.А.Лавров, А.Д.Тайманов, М.А.Тайцлин, Элементарные теории, Успехи математических наук 20 (1965) № 4, с.37-108.
- [ 3 ] F.Montagna, A completeness result for fixed-point algebras, Zeitschrift für mathematische Logik und Grundlagen der Mathematik 30 (1984), S.525-532.
- [ 4 ] C.Smoryński, Fixed point algebras, Bulletin (New Series) of the American Mathematical Society 6 (1982), pp.317-356.
- [ 5 ] C.Smoryński, The finite inseparability of the first-order theory of diagonalisable algebras, Studia Logica 41 (1982), pp.347-349.
- [ 6 ] C.Smoryński, Self-Reference and Modal Logic, Springer-Verlag, Berlin, 1985.
- [ 7 ] R.M.Solovay, Infinite fixed-point algebras, in: A.Nerode and R.A.Shore, eds., Recursion Theory, Proceedings of Symposia in Pure Mathematics, vol.42, AMS, Providence, 1985, pp.473-486.
- [ 8 ] М.А.Тайцлин, Эффективная неотделимость множества тождественно истинных и множества конечно опровержимых формул элементарной теории структур, Алгебра и логика 1 (1962) № 3, с.24-38.



# 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

## *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 Zambella  
*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 Gavaldà, 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

A Generalized Quantifier Logic for Naked Infinitives  
Dynamic Montague Grammar  
Concept Formation and Concept Composition  
Intuitionistic Categorical 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

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

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  $\text{I}\Delta_0 + \Omega_1$ , revised version

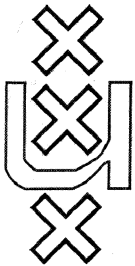
**Institute for Language, Logic and Information**

**ON THE PROOF OF  
SOLOVAY'S THEOREM**

Dick de Jongh  
Marc Jumelet  
Franco Montagna

ITLI Prepublication Series  
for Mathematical Logic and Foundations ML-89-04

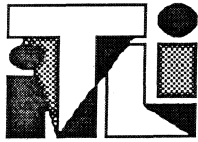
VOORBEELD



**University of Amsterdam**



hier oder ~~best~~



**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

# ON THE PROOF OF SOLOVAY'S THEOREM

Dick de Jongh  
Department of Mathematics and Computer Science, University of Amsterdam

Marc Jumelet  
Franco Montagna  
Dipartimento di Matematica, Università di Siena



## Introduction.

Solovay's arithmetical completeness theorem states that Löb's logic  $L$  (PRL in Smoryński[85]) is the modal logic of provability in  $PA$  and that the closure of  $L$  under reflection,  $\Box A \rightarrow A$ , and modus ponens is the provability logic of  $PA$  in the standard model. For any sentence  $\varphi$  such that  $L \not\vdash \varphi$ , Solovay defines, using the formalized recursion theorem, a recursive function from which an interpretation  $(\cdot)^*$  is obtained such that  $PA \not\vdash \varphi^*$ . The proofs of the essential properties of this function, as well as the formalization of the recursion theorem, employ, prima facie,  $\Sigma_1$ -induction.

In this article<sup>1</sup> we take another look at Solovay's proof of his completeness theorem for the modal logic  $L$  with respect to arithmetical interpretations. An at first sight dominant feature in Solovay's proof is his use of the formalized recursion theorem. The use of the recursion theorem in this proof and others like it is not really necessary, but can be replaced by applications of Gödel's diagonalization lemma (mostly in the form including free variables). Using the recursion theorem makes his procedure somewhat easier to follow intuitively, but it adds to the mystery of the proof, and makes it harder to judge exactly which principles are used. Since one of our purposes is to investigate in how far one can weaken the arithmetical system and still have Solovay's completeness result, it is important to us to do without it. The concrete additional benefits of the proof of the arithmetic completeness of  $L$  given in section 2 are:

- (1) it mainly uses modal properties of arithmetic as well as self-reference and is, therefore, closer to the spirit of modal logic;
- (2) the modal properties used, i.e. these of Guaspari-Solovay's  $R$  plus diagonalization are valid in weak fragments of  $PA$ ; they hold, for instance in any extension of  $I\Delta_0$  which proves  $\Sigma_1$ -completeness, so they hold e.g. in  $I\Delta_0 + EXP$ , but not in  $I\Delta_0 + \Omega_1$  (cf. Ver-

---

<sup>1</sup> Part of this article is a reworked version of the first chapter of the master's thesis of the second author (Jumelet [88])

brugge [88]). Consequently, the present proof allows us to extend Solovay's theorem to a large class of fragments of PA. The result concerning the provability logic of true formulae for these fragments falls under the scope of this proof as well.

The fixed point formulas used in the completeness proof for  $L$  are then, in section 3 of this paper, slightly modified to obtain a  $\Delta_0$ -formula describing the behaviour of Solovay's function. This formula is used to introduce, by means of the diagonalization lemma again, standard proof predicates provably equivalent to the usual one, yielding the arithmetical completeness of Guaspari and Solovay's system  $R$  with respect to extensions of  $I\Delta_0 + EXP$ .

## 1. Preliminaries.

1.1. Definition. The language  $L_{\Box}$  of propositional modal logic is defined as follows:

$L_{\Box} := \{\perp, \rightarrow, \Box, \neg, \wedge, \vee, \leftrightarrow\} \cup P$ , where  $P$  is some set of propositional letters,  $\perp$  a propositional constant (*falsum*),  $\rightarrow$  a binary connective (*material implication*) and  $\Box$  a modal operator. The class of well-formed formulae  $SEN_{L_{\Box}}$  of  $L_{\Box}$  is the smallest class such that:

$$\begin{aligned} P &\subseteq SEN_{L_{\Box}}, \\ \perp &\in SEN_{L_{\Box}}, \\ \varphi, \psi \in SEN_{L_{\Box}} &\Rightarrow (\varphi \rightarrow \psi) \in SEN_{L_{\Box}}, \\ \text{and } \varphi \in SEN_{L_{\Box}} &\Rightarrow \Box\varphi \in SEN_{L_{\Box}}. \end{aligned}$$

Boolean connectives  $\vee, \wedge, \neg, \leftrightarrow$ , as well as  $\Diamond$  will be used as abbreviations with their standard meaning.

1.2. Definition. A semantics for modal formulae is developed by means of so-called *Kripke-models*. A *model*  $M$  for  $L_{\Box}$  is a triple  $\langle M, R, \Vdash \rangle$ , where  $M$  is a non-empty set,  $R$  a binary relation on  $M$  and  $\Vdash$  some subset of  $M \times P$ .  $F = \langle M, R \rangle$  is called the *frame* of the model. The forcing relation is uniquely extended to all modal formulae  $\chi$  in the following manner (writing  $x \Vdash \chi$  for  $\langle x, \chi \rangle \in \Vdash$  and  $x \not\Vdash \chi$  for  $\langle x, \chi \rangle \notin \Vdash$ ):

for all  $x \in M$ :

for  $\chi = \varphi \rightarrow \psi$ :  $x \Vdash \chi$  iff  $x \nVdash \varphi$  or  $x \Vdash \psi$ ,

for  $\chi = \Box \varphi$ :  $x \Vdash \chi$  iff for all  $y \in M$  such that  $xRy$ :  $y \Vdash \varphi$ ,

and, finally,  $x \nVdash \perp$ .

**1.3. Definition.** The modal system that primarily concerns us here, is the so-called modal *provability logic* L. This system is defined as the smallest set of modal formulae containing:

all tautologies of propositional logic;

all expressions of the forms

$\Box \varphi \rightarrow \Box \Box \varphi$ ,  $\Box(\varphi \rightarrow \psi) \rightarrow (\Box \varphi \rightarrow \Box \psi)$ , or  $\Box(\Box \varphi \rightarrow \varphi) \rightarrow \Box \varphi$ ,

which is closed under the following two rules of inference:

$\vdash \varphi \Rightarrow \vdash \Box \varphi$  (*necessitation*);

$\vdash \varphi \rightarrow \psi$  and  $\vdash \varphi \Rightarrow \vdash \psi$ .

The axiom  $\Box \varphi \rightarrow \Box \Box \varphi$  is put on the list rather to stress its importance than its indispensability, since it can actually be derived from the other axioms and rules. The next result is of essential interest to us here.

**1.4. Theorem.**  $\varphi$  is not a theorem of L if and only if a model  $M := \langle M, R, \Vdash \rangle$  exists such that:

(i) M is finite, say  $M = \{1, \dots, n\}$ ;

(ii) R is a transitive and conversely well-founded relation on M, (i.e.:  $\forall x, y, z \in M (xRy \wedge yRz \rightarrow xRz)$  and no infinite ascending chain  $x_0 R x_1 R x_2 \dots$  of elements of M exists);

(iii) for all  $j \in M$ , if  $1 < j \leq n$ , then  $1Rj$ ;

(iv)  $1 \Vdash \neg \varphi$ .

This theorem is known as the *modal completeness theorem for L* with respect to the finite, transitive and conversely well-founded frames. For its proof one may consult e.g. Smoryński[85].

**1.5. Interpretations.** Let from now on T be a  $\Sigma_1$ -sound arithmetical theory proving the three Löb conditions (and hence Löb's theorem) and satisfying formalized  $\Sigma_1$ -completeness, i.e.:

$T \vdash \exists p \text{ proof}_T(p, \ulcorner A \rightarrow B \urcorner) \rightarrow (\exists p \text{ proof}_T(p, \ulcorner A \urcorner) \rightarrow \exists q \text{ proof}_T(q, \ulcorner B \urcorner))$ ;

$T \vdash A \rightarrow \exists p \text{ proof}_T(p, \ulcorner A \urcorner)$ , for all  $A \in \Sigma_1$ ;

An *interpretation* of a set of modal formulae is a function  $( )^*$  that assigns a sentence  $\varphi^*$  in the language of  $T$  to each modal expression  $\varphi$  and obeys the following criteria:

$$\begin{aligned} (\perp)^* &= 0 = 1; \\ (\varphi \rightarrow \psi)^* &= \varphi^* \rightarrow \psi^*; \\ (\Box\varphi)^* &= \exists p \text{ proof}_T(p, \ulcorner \varphi^* \urcorner). \end{aligned}$$

It is obvious that, once  $( )^*$  has been defined for each propositional variable in the modal language used, the translation of the entire set of formulae is completely determined.

### 1.6. Solovay's first Completeness Theorem (Solovay[76]).

This theorem is formulated as follows:

Let  $\varphi$  be any modal expression. Then:  $\vdash_L \varphi$  if and only if  $T \vdash \varphi^*$  for every interpretation  $( )^*$  of the modal language used which satisfies the clauses of the preceding paragraph.

The implication from the left to the right is of no concern to us here. The proof is simple, due to the fact that  $T$  is closed under the axioms and rules of  $L$  whenever the provability predicate is substituted for the modal operator  $\Box$ . The arithmetical versions of the rules and axioms of  $L$  are exactly the three Löb conditions and Löb's theorem which are fulfilled in  $T$ . The conditions imposed upon the interpretation function will do the rest. The implication in the other direction will be treated in section 2.

## 2. A modification of Solovay's completeness proof.

The original proof of the completeness theorem is based on the idea that a certain class of Kripke-models can be embedded in arithmetic. We have already seen that any modal expression  $\varphi$  which is not derivable from the axioms of  $L$  gives rise to some finite counter model falsifying  $\varphi$ . The embedding of such a model into arithmetic was carried out by Solovay by defining, with the aid of the recursion theorem, a recursive function  $h$  which paces through the model in a very particular way. Intuitively speaking, one can describe the Solovay function as follows. As its values it

takes only numbers denoting the nodes of the Kripke-model in question. The next value can only be the same as the previous one or one which is accessible from it by way of the relation  $R$  of the model. Thus it is clear that this function eventually reaches a limit. This limit is used to specify the next value, each time, in the following manner: for each argument the function takes the same value  $m$  as the previous one, unless the argument codes a proof in  $T$  of the fact that, for a certain number  $n$ ,  $R$ -accessible from  $m$ , the limit of the function is not equal to  $n$ . In the latter case the function takes this value  $n$ .

To be able to be more precise we now first give some notation.

**2.1. Definition.** Let  $F = \langle M, R \rangle$  be a finite, transitive and conversely well-founded frame.  $M = \{1, \dots, n\}$  and for all  $j$ , if  $1 < j \leq n$ , then  $1Rj$ . A new root  $0$  is added to  $M$ , i.e., for all  $j \in M$ ,  $0Rj$ .

We will use the following abbreviations:

$iRj$  for  $i=j \vee iRj$ ;

$i \not R j$  for  $\neg iRj \wedge \neg jRi$ .

The function  $h$  is represented by a formula  $Hxy$ . We write  $\ell = i$  for  $\exists x \forall y \geq x Hyi$ , i.e. "the limit of  $h$  is  $i$ ".

More formally, the function  $h$ , given by the formula  $Hxy$ , is defined as follows, using the formalized recursion theorem:

$h(0) = 0$

$h(n+1) = h(n)$  unless

$h(n)Rm$  and  $\text{proof}(n, \ulcorner \neg \ell = m \urcorner)$  in which case

$h(n+1) = m$ .

If the theory  $T$  is strong enough to allow definition by primitive recursion, the use of the recursion theorem can immediately be circumvented as follows. Let  $\text{nonlim}(u, v)$  be the function that, for each  $u$  and  $v$ , if  $u$  is the code of a formula  $Hxy$ , gives the code of  $\neg \exists x \forall y \geq x Hyv$ . One can then define  $h'(u, \cdot)$ , dependent on the extra variable  $u$ , simply by primitive recursion:

$h'(u, 0) = 0$

$h'(u, n+1) = h'(u, n)$  unless

$h'(u, n)Rm$  and  $\text{proof}(n, \text{nonlim}(u, m))$  in which case

$h'(u, n+1) = m$ .



If  $h'(u, x)$  is defined by  $H'uxy$ , then  $Hxy$  with properties as required can be found with the aid of the diagonalization lemma:

$$\vdash Hxy \leftrightarrow H'(\ulcorner Hxy \urcorner, x, y)$$

However, we do not want to have to depend on our theory to be strong enough to have primitive recursion available: in essence this still requires  $\Sigma_1$ -induction and it turns out that with definitions like the one given above there is no necessity for this. For the definition of  $h(n+1)$  we only have to look at numbers  $\leq n$  and the proofs of negations of limit assertions about  $h$  which they code.

Let us first consider the case of defining  $h$  only as a partial function at those arguments where relevant negations of limit assertions are actually proved. Then we can see that  $h(n+1)=m$  iff

- (1)  $n+1$  proves the negation of the limit assertion with respect to  $m$ ,
- (2) no such proof concerning a number  $m'$  with  $m R m'$  (or  $m=m'$ ) is coded by a number  $\leq n$  [otherwise,  $h$  should have "passed"  $m$  already],
- (3) if any such proof is coded by a number  $n' \leq n$  for an  $m'$  incomparable to  $m$  with respect to  $R$ , then there has to be an even smaller number  $n'' \leq n'$  that codes such a proof for a number  $m'' R m$  (or  $m''=m$ ) incomparable to  $m'$  [otherwise  $h$  should have taken a direction from which it could no longer reach  $m$ ; in other words, any proof that could possibly "side-track"  $h$  from its way to  $m$ , has to have been preceded by a proof that makes it harmless, by side-tracking it].

More formally a partial function can be thus defined as  $H_p$ , slightly changing the definition of  $\lambda=y$  to

$$\exists x (H_p xy \wedge \forall x' \geq x \neg \exists y' \leq n H_p x'y')$$

$$H_p xy \leftrightarrow (x=0 \wedge y=0) \vee$$

$$(\text{Proof}(x, \ulcorner \neg \lambda=y \urcorner) \wedge$$

$$\neg \exists x'' < x \exists y'' \leq n (y'' R y \wedge \text{Proof}(x'', \ulcorner \neg \lambda=y'' \urcorner)) \wedge$$

$$\forall x'' < x \forall y'' \leq n (y'' \circ y \wedge \text{Proof}(x'', \ulcorner \neg \lambda=y'' \urcorner) \rightarrow$$

$$\exists x''' < x'' \exists y''' (y''' \circ y'' \wedge y''' R y \wedge \text{Proof}(x''', \ulcorner \neg \lambda=y''' \urcorner))$$

$Hxy$  can then be obtained from  $H_p xy$  as follows:

$$Hxy \leftrightarrow \exists x' \leq x (H_p x'y \wedge \forall x'' (x' < x'' \leq x \rightarrow \neg \exists y' \leq n H_p x''y'))$$

This method of giving these definitions applies quite generally, and we will use it in section 3, but for the Solovay proof for L it can be further simplified. The proof involves only the mutual relations between a finite number of limit assertions, and we can more directly define corresponding sentences, using nothing but the desired connection between these sentences. More precisely, we may replace each expression " $\ell=i$ " we come across in the original proof, by a single sentence  $\lambda_i$ , the definition of which is an exact imitation of the conditions which lead to  $\ell=i$ . It is important to notice that these conditions can all be spelled out in the form of finite conjunctions, claiming the existence or non-existence and order of succession of certain proofs, namely proofs of expressions of the form  $\neg \ell=j$ . But within proof predicates only codes of these expressions occur. It turns out to be possible for that reason to define each  $\lambda_i$  by means of a fixed point equation, containing only codes of these  $\lambda_j$ 's. It will be demonstrated below, that, in doing so, the alternative sentences satisfy the same lemmas Solovay proved for the original ones. This makes them equally suitable to perform as a basis for arithmetical interpretations of the modal logic.

The  $n$ -ary fixed point theorem produces a set of sentences  $\lambda_0, \dots, \lambda_n$  in the language of  $T$ , which satisfy the following requirements:

$$T \vdash \lambda_1 \leftrightarrow \Box \neg \lambda_1 \wedge \bigwedge_{1 \leq i} \neg \Box \neg \lambda_i;$$

for all  $i$  such that  $1 < i \leq n$ :

$$T \vdash \lambda_i \leftrightarrow \Box \neg \lambda_i \wedge \bigwedge_{1 \leq j} \neg \Box \neg \lambda_j \wedge \bigwedge_{\substack{[p] \\ k \leq i \\ k \leq j}} (\Box \neg \lambda_k < \Box \neg \lambda_j).$$

Here " $\Box A < \Box B$ " is the usual notation for:

$$" \exists p [ \text{proof}_T(p, \ulcorner A \urcorner) \wedge \neg \exists q \leq p \text{proof}_T(q, \ulcorner B \urcorner) ]".$$

Finally, we define:

$$\lambda_0 := \neg \bigwedge_{1 \leq i \leq n} \lambda_i.$$

**2.2. Lemma.** The set of sentences  $\{\lambda_0, \dots, \lambda_n\}$  of  $T$  defined as above has the following properties:

- (1)  $T \vdash \bigvee_{0 \leq i \leq n} \lambda_i$ .
- (2)  $\mathbb{N} \vDash \lambda_0$ .
- (3) For all  $i$  such that  $0 \leq i \leq n$ ,  $T + \lambda_i$  is consistent.
- (4)  $T \vdash \lambda_i \rightarrow \bigwedge_{iRj} \neg \Box \neg \lambda_j$  for all  $i \geq 0$ .
- (5)  $T \vdash \lambda_i \rightarrow \bigwedge_{\neg iRj} \Box \neg \lambda_j$  for all  $i > 0$ .

This lemma represents the heart of Solovay's proof. If we replace each expression of the form  $\lambda_i$  by  $\mathfrak{L} = i$ , we get the original lemma (cf. Solovay[76], lemma 4.1).

For reasons of economy, it is useful to prove lemma 2.2 within a more general framework. This will show us exactly which properties of our theory are used to prove it. We take for this purpose a modified version of  $R^-$ , the modal system of Guaspari and Solovay (cf. Guaspari and Solovay[79]). We first recall that  $R^-$  is an extension of  $L$  in which the class of well-formed formulae is extended by the so-called witness comparison formulae, viz. those of the forms  $\Box A \prec \Box B$  and  $\Box A \preceq \Box B$ .

**2.3. Axioms of  $R^-$ .**  $R^-$  is axiomatized by adding to  $L$  the following axiom schemata (cf. de Jongh[87]):

$A \rightarrow \Box A$  for all boxed and witness comparison formulae. It is to be noted, that, since  $R^-$  is an extension of  $L$ , the same schema applies to the closure of this class under conjunctions and disjunctions, the so-called  $\Sigma$ -formulae; this gives us the so-called  $\Sigma$ -completeness axiom;

the *order axioms* (for all  $\Box$ -formulae  $A, B, C$ ):

- (01)  $A \rightarrow A \preceq B \vee B \preceq A$ ;
- (02)  $A \preceq B \rightarrow A$ ;
- (03)  $A \preceq B \wedge B \preceq C \rightarrow A \preceq C$ ;
- (04)  $A \prec B \leftrightarrow A \preceq B \wedge \neg B \preceq A$ .

We extend  $R^-$  as follows: for any frame  $F = \langle M, R \rangle$ , which is finite, transitive and conversely well-founded, with  $M = \{1, \dots, n\}$  and  $1Ri$  for all  $i$  such that  $1 < i \leq n$ , let  $R_F^-$  be defined by the addition of the following axioms to  $R^-$  (we assume the language to contain propositional constants  $L_0, \dots, L_n$ , and we write  $\Box A$  for  $A \wedge \Box A$ ):

$$\Box (L_1 \leftrightarrow \Box \neg L_1 \wedge \bigwedge_{1Ri} \neg \Box \neg L_i);$$

for each  $i$  such that  $1 < i \leq n$ :

$$\Box (L_i \leftrightarrow \Box \neg L_i \wedge \bigwedge_{iRj} \neg \Box \neg L_j \wedge \bigwedge_{i \circ j} \bigvee_{\substack{kBj \\ k \circ j}} (\Box \neg L_k \prec \Box \neg L_j));$$

$$\Box (L_0 \leftrightarrow \neg \bigvee_{1 \leq i \leq n} L_i).$$

These axioms will be referred to as the *limit axioms*. In addition, we let  $R_F^-$  contain

$$\Box (\neg (\Box \neg L_i \preceq \Box \neg L_j \wedge \Box \neg L_j \preceq \Box \neg L_i))$$

for all  $i, j$  such that  $0 \leq i, j \leq n$  and  $i \neq j$ , as so-called *proof apartness axioms*. In the next two paragraphs we will mention some properties of  $R_F^-$  that will be needed for the proof of lemma 2.2.

In the following discussion the frame  $F$  is to be thought of as fixed.

**2.4. Theorem (Soundness of  $R_F^-$ ).** An interpretation  $( )^+$  of sentences in the language of  $R_F^-$  into the language of arithmetic is called *F-sound* if and only if  $( )^+$  fulfils the criteria cited for  $( )^*$  in 1.5 and, in addition:

for all formulae  $\varphi, \psi$ :

$$(\Box \varphi \preceq \Box \psi)^+ = \exists p [\text{proof}_T(p, \ulcorner \varphi^+ \urcorner) \wedge \neg \exists q < p \text{proof}_T(q, \ulcorner \psi^+ \urcorner)];$$

$$(\Box \varphi \prec \Box \psi)^+ = \exists p [\text{proof}_T(p, \ulcorner \varphi^+ \urcorner) \wedge \neg \exists q \leq p \text{proof}_T(q, \ulcorner \psi^+ \urcorner)];$$

for all  $i$  such that  $0 \leq i \leq n$ :

$$L_i^+ = \lambda_i \text{ (as defined above).}$$

For all *F-sound* interpretations  $( )^+$  of sentences in the language of  $R_F^-$  and any  $\varphi$  in that language,  $R_F^- \vdash \varphi \Rightarrow T \vdash \varphi^+$ .

The proof is straightforward by induction on the length of proof in  $R_F^-$ , since  $T$  is closed under the same rules and axioms we have at our disposal in  $R_F^-$ , provided  $( )^+$  is *F-sound*. We will use this theorem extensively in the proof of lemma 2.2.

A Kripke-model for  $R^-$  is a finite, tree-ordered Kripke-model  $\langle X, U, \Vdash \rangle$  for  $L$  in which witness-comparison formulae are treated as if they were atomic formulae and in which the following requirements are fulfilled:

(1) *persistence* of  $\prec$  and  $\preceq$ :

if  $s \Vdash A \preceq B$  and  $s U s'$ , then  $s' \Vdash A \preceq B$ ,  
and likewise for  $\prec$ , viz.:

if  $s \Vdash A \prec B$  and  $s U s'$ , then  $s' \Vdash A \prec B$ ;

(2) each instance of the order axioms is satisfied at each node.

The completeness theorem for  $R^-$  is stated as follows:  $R^- \vdash \varphi$  iff  $\varphi$  is valid on all finite, tree-ordered Kripke-models for  $R^-$ .

In the case of  $R_F^-$  this theorem implies:

### 2.5. Theorem (*completeness of $R_F^-$* ).

If  $R_F^- \not\vdash \varphi$ , then a finite, tree-ordered Kripke-model for  $R^-$  exists, in which all limit axioms and proof apartness axioms are forced at each node, and on which  $\varphi$  is falsified.

*Proof.* This result is a consequence of the completeness theorem for  $R^-$ , because we have :

$$R_F^- \vdash \varphi \iff R^- \vdash \theta \rightarrow \varphi,$$

where  $\theta$  is the finite conjunction of limit axioms and proof apartness axioms listed in the definition of  $R_F^-$ .

The implication from the right to the left is easily proved. The other direction is shown by induction on the length of the proof in  $R_F^-$ . To obtain the desired result, we should check whether any proof of a formula  $\varphi$  in  $R_F^-$  can be transformed into a proof of  $\theta \rightarrow \varphi$  in  $R^-$ . This can cause no difficulty, since any axiom of  $R_F^-$  is either an axiom of  $R^-$  or a consequence of  $\theta$ , and, if the last rule applied in a proof in  $R_F^-$  of some formula  $\psi$  has been the necessitation rule, then we can use  $\theta \rightarrow \Box \theta$  which is a theorem of  $R^-$ .  $\square$

A simple proof of the completeness theorem for  $R^-$  can be found in De Jongh [87].

Now we are ready to commence the proof of lemma 2.2.

Proof of lemma 2.2.

Fix a finite, transitive and conversely well-founded frame  $F = \langle M, R \rangle$ , with  $M = \{1, \dots, n\}$  and  $1Ri$  for all  $i$  such that  $1 < i \leq n$ . Let  $\lambda_0, \dots, \lambda_n$  and  $R_F^-$  be as defined above. We first show:

$$(a) R_F^- \vdash L_0 \leftrightarrow \bigwedge_{1 \leq i \leq n} \Box \neg L_i.$$

As the implication from the right to the left is obvious, we will concentrate on the opposite direction. Suppose the contrary to be the case. We will derive a contradiction as follows. By theorem 2.5 we would have a finite, tree-ordered Kripke-model  $\langle X, U, \Vdash \rangle$  for  $R^-$  with the limit and proof apartness axioms forced everywhere in the model and with some bottom node  $k_0$  such that

$$k_0 \Vdash L_0 \wedge \bigvee_{1 \leq i \leq n} \Box \neg L_i.$$

We must have:

$$k_0 \Vdash \Box \neg L_{i_1} \wedge \dots \wedge \Box \neg L_{i_k} \wedge \bigwedge_{j \notin \{i_1, \dots, i_k\}} \Box \neg L_j,$$

for some  $k$  such that  $1 \leq k \leq n$ .

As any instance of the order axioms and the proof apartness axioms is forced at  $k_0$ , we can stipulate, without loss of generality, that at  $k_0$  the following is forced:

$$\Box \neg L_{i_1} < \Box \neg L_{i_2} \wedge \dots \wedge \Box \neg L_{i_{k-1}} < \Box \neg L_{i_k}.$$

At this point, we can construct a subset  $\{m_1, \dots, m_l\}$  of the set of indices  $\{1, \dots, k\}$  as follows:

$$m_1 := 1;$$

$$m_{h+1} := m \text{ if } m \text{ is the smallest index number in } \{1, \dots, k\} \text{ such}$$

$$i_{m_h} R i_m \text{ and } k_0 \Vdash \Box \neg L_{i_{m_h}} < \Box \neg L_{i_m}. \text{ If no such } m \text{ exists, set } l = h \text{ and } m_{h+1} = m_h.$$

It will be understood that this construction comes to an end, because the set  $\{1, \dots, k\}$  is finite. By means of a finite induction procedure we will now prove the following: for all  $p$  such that  $1 \leq p \leq l$ :

$$k_0 \Vdash \bigwedge_{j \in \text{omp}} \bigvee_{\substack{k \in j \\ k R m_p}} (\Box \neg L_k < \Box \neg L_j).$$

The case of  $p = 1$  is trivial, since  $i_{m_1} = i_1$ .

Induction step: suppose

$$k_0 \Vdash \bigwedge_{j \in m_p} \bigvee_{\substack{koj \\ k \in Bm_p}} (\Box \neg L_k \prec \Box \neg L_j).$$

Now let  $j$  be such, that  $j \in m_{p+1}$ . There are two possibilities: either  $j \in m_p$  as well, or not. In the first case we obtain

$$k_0 \Vdash \bigvee_{\substack{koj \\ k \in Bm_{p+1}}} (\Box \neg L_k \prec \Box \neg L_j)$$

by the induction hypothesis, for  $k \in Bm_p$  implies  $k \in Bm_{p+1}$ . In the latter case  $m_p R j$  must hold. But the definition of  $m_{p+1}$  implies:  $k_0 \Vdash \Box \neg L_{i_{m_{p+1}}} \prec \Box \neg L_j$  whence  $k_0 \Vdash \bigvee_{\substack{koj \\ k \in Bm_{p+1}}} (\Box \neg L_k \prec \Box \neg L_j)$  follows by propositional logic.

This completes the induction procedure. Since  $i_{m_1}$  has no  $R$ -successors in  $\{i_1, \dots, i_k\}$ , we can conclude:

$$k_0 \Vdash \Box \neg L_{i_{m_1}} \wedge \bigwedge_{i_{m_1} R j} \neg \Box \neg L_j \wedge \bigwedge_{j \in i_{m_1}} \bigvee_{\substack{koj \\ i_{m_1} B j}} (\Box \neg L_k \prec \Box \neg L_j).$$

But this implies  $k_0 \Vdash L_{i_{m_1}}$  contradicting  $k_0 \Vdash L_0$ . The proof is hereby completed.

(b) If  $1 \leq i \leq n$ , then  $R_F^- \vdash L_i \rightarrow \bigwedge_{i R j} \neg \Box \neg L_j$ . This is immediate from the definition of  $R_F^-$ .

Combining (a) and (b) we get lemma 2.2(4) by soundness.

(c)  $R_F^-$  contains all tautologies of propositional logic, so we have  $R_F^- \vdash L_0 \vee \neg L_0$  from which  $R_F^- \vdash \bigvee_{0 \leq i \leq n} L_i$  readily follows. Employing soundness, this accounts for of lemma 2.2(1).

As all theorems of  $T$  hold in the standard model, we must have  $\mathbb{N} \models \lambda_i$  for some  $i$  such that  $0 \leq i \leq n$ . But it must be the case that  $\mathbb{N} \models \lambda_0$ , since for any  $i \neq 0$  we would have  $T \vdash \neg \lambda_i$  in case  $\lambda_i$  were true. Combining this with of lemma 2.2(4), we obtain

$$\mathbb{N} \models \bigwedge_{0 \leq i \leq n} \neg \Box \neg \lambda_j. \text{ This settles lemma 2.2(2) and (3).}$$

(d) If  $0 < i \leq n$ , then  $R_F^- \vdash L_i \rightarrow \Box \neg L_0$ .

By (a) we have  $R_F^- \vdash \Box \neg L_i \rightarrow \neg L_0$ . Applying the necessitation rule we infer:  $R_F^- \vdash \Box \Box \neg L_i \rightarrow \Box \neg L_0$ . As  $\Box \neg L_i$  is a boxed formula,  $\Box \neg L_i \rightarrow \Box \Box \neg L_i$  is a theorem of  $R_F^-$ . This completes the proof, as  $R_F^- \vdash L_i \rightarrow \Box \neg L_i$  is a direct consequence of the definition of  $R_F^-$ .

(e) If  $0 < i \leq n$  and  $iRj$ , then  $R_F^- \vdash L_j \rightarrow \Box \neg L_i$ .

If  $iRj$  is the case, we have  $R_F^- \vdash \Box \neg L_j \rightarrow \neg L_i$  by the limit axiom that defines  $L_i$ . Arguing as in (d) we obtain the desired result.

(f) If  $0 < i \leq n$  and  $0 < j \leq n$  and  $i \circ j$ , then  $R_F^- \vdash L_i \rightarrow \Box \neg L_j$ .

Fix  $i$  and  $j$  such that  $i \circ j$ . By the definition of  $R_F^-$  we have:

$$R_F^- \vdash L_i \rightarrow \bigwedge_{i \circ j'} \bigvee_{\substack{kBi \\ k \circ j'}} (\Box \neg L_k \prec \Box \neg L_{j'}).$$

More specifically, we obtain:

$$R_F^- \vdash L_i \rightarrow \bigwedge_{\substack{i \circ j' \\ j'Bj}} \bigvee_{\substack{kBi \\ k \circ j'}} (\Box \neg L_k \prec \Box \neg L_{j'}).$$

As the order axioms and proof apartness axioms imply that the  $\Box \neg L_k$ 's in this formula are linearly ordered by  $\prec$  (compare the proof of (a)), there must be a smallest one; in other words:

$$R_F^- \vdash \bigwedge_{\substack{i \circ j' \\ j'Bj}} \bigvee_{\substack{kBi \\ k \circ j'}} (\Box \neg L_k \prec \Box \neg L_{j'}) \rightarrow \bigvee_{\substack{k \circ j \\ kBi}} \bigwedge_{\substack{j'Bj \\ j' \circ i}} (\Box \neg L_k \prec \Box \neg L_{j'}).$$

But the consequent in the last formula is a  $\Sigma$ -expression implying  $\neg L_j$ , so:  $R_F^- \vdash \bigwedge_{\substack{i \circ j' \\ j'Bj}} \bigvee_{\substack{kBi \\ k \circ j'}} (\Box \neg L_k \prec \Box \neg L_{j'}) \rightarrow \Box \neg L_j$ .

(g) Putting (d), (e) and (f) together, we obtain:

$$R_F^- \vdash L_i \rightarrow \Box (\neg L_0 \wedge \bigwedge_{i \circ j} \neg L_j \wedge \bigwedge_{jRi} \neg L_j) \text{ for all } i \text{ such that } 0 < i \leq n.$$

Applying soundness, this settles lemma 2.2 (5).  $\square$

Let  $M = \langle M, R, \Vdash \rangle$  be a finite, transitive and conversely well-founded model with  $M = \{1, \dots, n\}$  and for all  $i$  if  $1 < i \leq n$ , then  $1Ri$ . As usual, we expand  $M$  by adding an extra node  $0$  to it and defining  $0 \Vdash$  as equivalent to  $1 \Vdash$  for all propositional letters. In the manner indicated above we obtain sentences  $\lambda_0, \dots, \lambda_n$  satisfying lemma 2.2. We define an interpretation  $( )^*$  by setting for all  $p \in P$ :

$$p^* := \bigvee_{i \Vdash p} \lambda_i. \text{ If there is no } i \text{ such that } i \Vdash p, \text{ then set:} \\ p^* := "0 = 1"$$

The following lemma provides the necessary last step towards the completeness theorem:



2.7 Lemma: for all modal expressions  $\varphi$ , if  $1 \leq i \leq n$ , then:

$i \Vdash \varphi \Rightarrow T \vdash \lambda_i \rightarrow \varphi^*$  and

$i \nVdash \varphi \Rightarrow T \vdash \lambda_i \rightarrow \neg \varphi^*$ .

The proof is exactly the same as Solovay's original one, with each expression of the form  $\ell = i$  replaced by  $\lambda_i$ , so we will not give it here. Some attention however should be paid to the way clause (5) of lemma 2.2, in the form  $\lambda_i \rightarrow \Box \bigvee_{i \in R_j} \lambda_j$ , is used, when  $i \Vdash \varphi \Rightarrow T \vdash \lambda_i \rightarrow (\varphi)^*$  is proved by induction. In fact, it is at this point that full formalized  $\Sigma_1$ -completeness is used. This completes our explanation concerning the adaptation of the proof of Solovay's result.

### 3. Completeness of R.

In this section, we deal with the arithmetical completeness of Guaspari and Solovay's logic R with respect to arithmetical interpretations in  $I\Delta_0 + EXP$  or in any given  $\Sigma_1$ -sound RE-extension of it. To formulate our result correctly, let us start with the following definitions (as usual, T denotes an arbitrary  $\Sigma_1$ -sound RE-extension of  $I\Delta_0 + EXP$ ).

3.1. Definition. A *standard proof predicate* for T is a  $\Sigma_1$ -formula  $Th(v)$  numerating the set of theorems of T and such that for any two sentences  $\alpha, \beta$ ,  $T \vdash Th(\ulcorner \alpha \urcorner) \wedge Th(\ulcorner \alpha \rightarrow \beta \urcorner) \rightarrow Th(\ulcorner \beta \urcorner)$  and  $T \vdash Th(\ulcorner \alpha \urcorner) \rightarrow Th(\ulcorner Th(\ulcorner \alpha \urcorner) \urcorner)$

In our proof, we shall make use of a standard proof predicate which, in addition, is provably equivalent to the usual one.

3.2. Definition. Let  $Th(v)$  be a standard proof predicate for T. An *arithmetical interpretation based on  $Th(v)$*  is a mapping  $*$  from R formulas into arithmetical sentences satisfying the following conditions:  $\perp^* \equiv 0=1$ ,  $\top^* \equiv 0=0$ ;  $*$  commutes with the logical connectives and witness comparisons, and  $\Box A^* \equiv Th(\ulcorner A^* \urcorner)$ .

We are now ready to state the main theorem of this section.

**3.3. Theorem.** Let  $A$  be any formula of  $R$ . The following are equivalent:

- (i)  $R \vdash A$ .
- (ii) For any standard proof predicate  $Th(v)$  and for any interpretation  $*$  based on it,  $T \vdash A^*$ .
- (iii) For any standard proof predicate  $Th(v)$  provably equivalent to the usual one and for each interpretation based on it,  $T \vdash A^*$ .

Proof. That (i) implies (ii) is easy, and that (ii) implies (iii) is trivial. So, let us prove that (iii) implies (i). Suppose  $R \not\vdash A$ . By a result of Guaspari and Solovay (cf. [79]) there is a model  $M = \langle \{1, \dots, n\}, R, \Vdash \rangle$  of  $R^-$  with root 1 and a node  $i$  of  $M$  such that  $i \not\Vdash A$ ; moreover, the model can be taken to be  $A$ -sound, i.e. we can assume that  $1 \Vdash \Box B \rightarrow B$  for any subformula  $\Box B$  of  $A$ . Add a new node 0, stipulate that  $0 R i$  for  $i = 1, \dots, n$ , and give 0 the same forcing as 1 w.r.t. the subformulas of  $A$ . That this is possible is guaranteed by the fact that the model is  $A$ -sound.

Let  $S$  denote the set of  $\Box$ -subformulas of  $A$ ,  $K$  denote the cardinality of  $S$  plus one. For  $i = 0, \dots, n$  and for  $\Box C, \Box D \in S$  define:  $\Box C \equiv_i \Box D$  iff  $i \Vdash \Box C \leq \Box D$  and  $i \Vdash \Box D \leq \Box C$ ;  
 $\Box C <_i \Box D$  iff  $i \Vdash \Box C < \Box D$ . Furthermore, let  $E_{i_1}, \dots, E_{i_{h_i}}$  be the equivalence classes w.r.t.  $\equiv_i$  enumerated according to  $<_i$  (i.e. if  $\Box C \in E_{i_j}$ ,  $\Box D \in E_{i_h}$ , and  $j < h$ , then  $\Box C <_i \Box D$ ). Notice that, for  $i = 0, \dots, n$ ,  $h_i < K$ .

We add some more notation:

$\text{proof}(v, \ulcorner p \urcorner) := \text{proof}_T(v, \ulcorner p \urcorner)$ ;

$\Box \ulcorner p \urcorner := \exists v \text{ proof}(v, \ulcorner p \urcorner)$ ;

$\Box_{\leq x} \ulcorner p \urcorner := \exists v \leq x \text{ proof}(v, \ulcorner p \urcorner)$ ;

$\Box_{\leq x} \ulcorner p \urcorner \leq \Box_{\leq x} \ulcorner q \urcorner := \exists v \leq x [\text{proof}(v, \ulcorner p \urcorner) \wedge \forall u < v \neg \text{proof}(u, \ulcorner q \urcorner)]$ ;

$\Diamond_{\leq x} \ulcorner p \urcorner := \neg \Box_{\leq x} \ulcorner \neg p \urcorner$ ;

$\vdash := T \vdash$ .

**3.4. Definition.** A formula  $A$  is *stable* iff

$\vdash \exists x (\forall y \geq x Ay \vee \forall y \geq x \neg Ay)$ .

### 3.5. Lemma.

(1) Each Boolean combination of stable formulas in the same free variable  $x$  is a stable formula.

(2)  $\Box_{\leq x} \ulcorner p \urcorner$ ,  $\Diamond_{\leq x} \ulcorner p \urcorner$ ,  $\Box_{\leq x} \ulcorner p \urcorner \preceq \Box_{\leq x} \ulcorner q \urcorner$  are stable.

(3) if  $L(A_1(x), \dots, A_n(x))$  is a lattice combination of stable formulas  $A_1(x), \dots, A_n(x)$ , and if  $L_i \equiv \exists y \forall x \geq y A_i(x)$ , then:

$$\vdash \exists y \forall x \geq y L(A_1(x), \dots, A_n(x)) \leftrightarrow L(L_1, \dots, L_n).$$

Proof. (1) and (2) are trivial, and (3) is proved by induction on the complexity of  $L$ . The step corresponding to  $\wedge$  is trivial; the step corresponding to  $\vee$  is proved by means of (1) and the induction hypothesis.

Next let the free variable formulas  $H_i(x)$  for  $1 \leq i \leq n$  be defined, by self-reference, in such a way that:

$$\vdash H_i(x) \leftrightarrow \Box_{\leq x} \neg L_i \wedge \bigwedge_{iRj} \Diamond_{\leq x} L_j \wedge \bigwedge_{i \circ j} \bigvee_{\substack{kBi \\ k \circ j}} (\Box_{\leq x} \neg L_k \prec \Box_{\leq x} \neg L_j);$$

where  $iRj$  and  $i \circ j$  are defined as in 2.1 and  $L_i := \exists y \forall x \geq y H_i(x)$ .

Also, let  $H_0(x) := \bigwedge_{i \neq 0} \neg H_i(x)$ .

By lemma 3.2,  $H_i(x)$ ,  $i=0, \dots, n$ , are stable. Therefore, by the same lemma, clause (3):

$$\begin{aligned} \vdash L_0 &\leftrightarrow \bigwedge_{i \neq 0} \neg L_i \\ \vdash L_i &\leftrightarrow \Box_{\leq x} \neg L_i \wedge \bigwedge_{iRj} \neg \Box_{\leq x} \neg L_j \wedge \bigwedge_{i \circ j} \bigvee_{\substack{kBi \\ k \circ j}} (\Box_{\leq x} \neg L_k \prec \Box_{\leq x} \neg L_j); \end{aligned}$$

(of course we use:  $\vdash \Box \ulcorner p \urcorner \leftrightarrow \exists x \Box_{\leq x} \ulcorner p \urcorner$ ,  $\vdash \Diamond \ulcorner p \urcorner \leftrightarrow \forall x \Diamond_{\leq x} \ulcorner p \urcorner$ ,  $\vdash \Box \ulcorner p \urcorner \preceq \Box \ulcorner q \urcorner \leftrightarrow \exists x (\Box_{\leq x} \ulcorner p \urcorner \preceq \Box_{\leq x} \ulcorner q \urcorner)$ ).

As in lemma 2.2, we can now deduce:

$$(1) \vdash \bigvee_{0 \leq i \leq n} L_i.$$

$$(2) \not\vdash L_0.$$

(3) For all  $i$  such that  $0 \leq i \leq n$ ,  $T+L_i$  is consistent.

$$(4) \vdash L_i \rightarrow \bigwedge_{iRj} \neg \Box_{\leq x} \neg L_j \text{ for all } i \geq 0.$$

$$(5) \vdash L_i \rightarrow \bigwedge_{\neg iRj} \Box_{\leq x} \neg L_j \text{ for all } i > 0. \quad \square$$

### 3.6. Lemma.

(1) If  $i \neq j$ , then  $\vdash H(x, i) \rightarrow \neg H(x, j)$ ;

(2)  $\vdash H(x, i) \rightarrow \bigvee_{iBj} H(x+y, j)$ .

Proof.

(1) Suppose  $i \neq j$ . If  $i R j$ , then  $\vdash H(x, i) \rightarrow \Diamond_{\leq x} L_j$  and  $\vdash H(x, i) \rightarrow \Box_{\leq x} \neg L_j$ . The reasoning in the case  $j R i$  is symmetric. If  $i \circ j$ , one can formalize the following argument: assume  $H(x, i)$  and  $H(x, j)$ . Then, for each  $h$  incomparable with  $i$ , there is a  $k$  such that  $k \circ h$ ,  $k R i$  and  $\Box_{\leq x} \neg L_k \prec \Box_{\leq x} \neg L_h$ . Moreover, a similar condition holds with  $j$  in place of  $i$ . Since  $i \circ j$ ,  $H(x, i)$  implies that there is an  $h_1$  such that  $h_1 R i$ ,  $h_1 \circ j$  and  $\Box_{\leq x} \neg L_{h_1} \prec \Box_{\leq x} \neg L_j$ . Using  $H(x, j)$ , we get, since  $h_1 \circ j$ , an  $h_2 R j$ ,  $h_2 \circ h_1$  such that  $\Box_{\leq x} \neg L_{h_2} \prec \Box_{\leq x} \neg L_{h_1}$ . Thus,  $h_1, \dots, h_{n+1}$  are obtained such that  $\Box_{\leq x} \neg L_{h_{n+1}} \prec \Box_{\leq x} \neg L_{h_n} \prec \dots \prec \Box_{\leq x} \neg L_{h_1}$ . The proof apartness condition implies that the  $h_i$ 's are mutually distinct. This is impossible as  $X$  has cardinality  $n$ .

(2) Induction on  $y$  (notice that the formula  $H(x, i) \rightarrow \bigvee_{i R j} H(x+y, j)$  is  $\Delta_0$ ). Assume  $H(x+y, j)$ , where  $i R j$ . Clearly, if  $\neg \Box_{\leq x+y+1} \neg L_h$  for any  $h$  such that  $j R h$ , then  $H(x+y+1, j)$  and we are done; otherwise, let  $h$  be such that  $\Box_{\leq x+y+1} \neg L_h$  and  $j R h$ . Note that:

(a)  $\text{proof}(x+y+1, \neg L_h) \wedge \forall u < x+y+1 \neg \text{proof}(x+y+1, \neg L_h)$ ,

since otherwise we would have  $\neg H(x+y, j)$ ;

(b) if  $h R k$ , then  $\neg \Box_{\leq x+y+1} \neg L_k$ , otherwise, since  $\text{proof}(x+y+1, \neg L_h)$ , and, consequently  $\text{proof}(x+y+1, \neg L_k)$ , we would get  $\Box_{\leq x+y} \neg L_k$  and  $\neg H(x+y, j)$ ;

(c) if  $m \circ h$ , then, either  $j R m$ , in which case  $\neg \Box_{\leq x+y} \neg L_m$ ,  $\neg \Box_{\leq x+y+1} \neg L_m$ , and, since, by (a),  $\Box_{\leq x+y+1} \neg L_h$ , we can conclude  $\Box_{\leq x+y+1} \neg L_h \prec \Box_{\leq x+y+1} \neg L_m$ ,

or  $j \circ m$ , in which case there exists  $l$  such that  $l \circ m$ ,  $l R j$  (whence  $l R h$ ) and  $\Box_{\leq x+y} \neg L_h \prec \Box_{\leq x+y} \neg L_m$  (whence

$\Box_{\leq x+y+1} \neg L_l \prec \Box_{\leq x+y+1} \neg L_m$ ). In any case, if  $m \circ h$ , there exists  $l$

(possibly  $l=h$ ) such that  $l R h$ ,  $l \circ m$ , whence

$\Box_{\leq x+y+1} \neg L_l \prec \Box_{\leq x+y+1} \neg L_m$ . Conclusion:  $H(x+y+1, h)$ . This completes our proof.  $\square$

### 3.7. Corollary.

$\vdash H(x, i) \wedge y > x \rightarrow \bigvee_{i R j} H(y, j)$ .

We now introduce a standard proof predicate  $\Box' \ulcorner p \urcorner \equiv \exists x \text{proof}'(x, \ulcorner p \urcorner)$ , such that  $\top \not\vdash A^*$ , where  $*$  is the interpretation based on  $\Box'$  given by:

$$p_i^* := (i \equiv i) \wedge \bigwedge_{j \Vdash p_i} L_j$$

Roughly speaking,  $\square'$  proves  $p$  at stage  $Kx$  ( $K$  the cardinality of  $S$  plus one) iff  $\text{proof}(x, p)$  and, for all  $\square B \in S$ ,  $p \neq B^*$ , and proves  $p$  at stage  $Kx+y$  ( $0 < y < K$ ) iff  $\exists i \leq n [H(x, i) \wedge y \leq h_i \wedge \exists \square B \in E_{iy} (p = B^*)]$ . So, if  $H(x, i)$  and  $E_{iy} = \{\square B_1, \dots, \square B_s\}$ , then  $\square'$  proves  $B_1^*, \dots, B_s^*$  at stage  $Kx+y$ . Of course, the definition of  $\square'$  depends on the interpretation  $*$  which in fact is based on it. This circularity is avoided as usual by means of the diagonalization lemma. We will now present the formal definition of  $\text{proof}'$ :

### 3.8. Definition.

Let, by self-reference, the formula  $\text{proof}'(x, p)$  be such that:

$$\begin{aligned} \vdash \text{proof}'(x, p) \leftrightarrow & \exists y \leq x (Ky = x \wedge \text{proof}(y, p) \wedge \forall \square B \in S (\neg x = B^*)) \vee \\ & \exists i \leq n \exists y < x \exists z (0 < z < K \wedge x = Ky + z \wedge H(y, i) \wedge \\ & z \leq h_i \wedge \exists \square B \in E_z (p = B^*)) \end{aligned}$$

where  $*$  denotes the interpretation based on  $\exists x \text{proof}'(x, p)$  given by:  $p_i^* := (i \equiv i) \wedge \bigwedge_{j \Vdash p_i} L_j$ .

Notice that  $\text{proof}'$  is provably  $\Delta_0$ . To prove Theorem 3.3, it is sufficient to show (cf. Smoryński[85] or Guaspari-Solovay[79]) the following lemmas:

3.9. Lemma. If  $B$  is a subformula of  $A$ , then for all  $i \leq n$ :

$$\vdash L_i \rightarrow (B^* \leftrightarrow i \Vdash B).$$

3.10. Lemma. If  $\square B$  is a subformula of  $A$ , then for all  $i \leq n$ :

$$\vdash L_i \rightarrow (\square B^* \leftrightarrow i \Vdash \square B).$$

3.11. Lemma.  $\vdash \forall x (\square x \leftrightarrow \square' x)$ .

Proof of lemma 3.9. By induction on the complexity of  $B$ ; the proof works as in Guaspari-Solovay[79]. The only problem is that we have to be careful with the use of induction. But, even if we want to allow only  $\Delta_0$ -induction, there can be no problem, since both  $H(x, y)$  and  $\text{proof}'$  are  $\Delta_0$ . Anyway, the propositional cases and the Boolean cases are trivial.

$\Box$ -case: if  $\Box B$  is a subformula of  $A$ , then  $B^*$  can be proved only at a stage of the form  $Kx+z$ , where  $0 < z < K$ . This happens iff  $H(x,i)$  and  $i \Vdash \Box B$ ; so:  $\vdash (L_1 \wedge i \Vdash \Box B) \rightarrow (\Box B)^*$ .

Next, suppose  $L_1 \wedge i \Vdash \neg \Box B$ . Lemma 3.6 and corollary 3.7 and the definition of  $H(x,0)$  ensure that, provably in  $T$ ,  $H(x,y)$  defines the graph of a weakly monotonic function from  $\mathbb{N}$  to  $\{0, \dots, n\}$ . So,  $L_1 \wedge i \Vdash \neg \Box B$  implies that, for all  $x$ :

$$\begin{aligned} H(x,j) &\rightarrow j \Vdash i \\ &\rightarrow j \nVdash \Box B. \end{aligned}$$

So,  $B^*$  is never proved by proof'.

Steps  $\prec, \preceq$ . Suppose  $L_1 \wedge i \Vdash \Box B \prec \Box C$ ; then there is a least  $x$  such that:  $\exists j \leq n [H(x,j) \wedge j \Vdash \Box B]$  (we have applied the least number principle to the  $\Delta_0$ -formula  $\exists j \leq n [H(x,j) \wedge j \Vdash \Box B]$ ). Note that, by lemma 3.6, this  $j$  is unique and  $j \Vdash i$ , and therefore, by  $\Sigma$ -persistence,  $j \Vdash \Box B \prec \Box C$ .

If  $u < x$ , then:  $H(u,h) \rightarrow h \nVdash \Box C$  (otherwise,  $h R j$ ,  $h \Vdash \Box C \prec \Box B$ ,  $j \Vdash \Box C \prec \Box B$ ). So,  $C^*$  is not proved by proof' at any stage  $\leq Kx$ . Since  $j \Vdash \Box B \prec \Box C$ , we get either  $j \Vdash \neg \Box C$  or  $\Box B \in E_{j_r}$ ,  $\Box C \in E_{j_s}$  where  $r < s$ . It follows, that  $B^*$  is proved at stage  $Kx+r$ , and either  $C^*$  is proved at stage  $Kx+s$ ,  $s > r$  or  $\neg \Box_{\leq K(x+1)} C^*$ . In both cases,  $(\Box B)^* \prec (\Box C)^*$ .

The case  $L_1 \wedge i \Vdash \Box B \preceq \Box C$  is treated similarly.

If  $L_1 \wedge i \nVdash \Box B \prec \Box C$ , then either  $i \nVdash \Box B$  and  $\neg (\Box B)^*$  by the  $\Box$ -step, or it is the case that  $i \Vdash \Box C \preceq \Box B$ , in which case  $(\Box C \preceq \Box B)^*$ , whence  $\neg (\Box B \prec \Box C)^*$  follows.

The case  $L_1 \wedge i \nVdash \Box B \preceq \Box C$  is treated similarly.

Proof of lemma 3.10. By conditions (1), ..., (5) of lemma 2.2 and by lemma 3.9, we are in a position to repeat the proof of the analogous lemma in Guaspari-Solovay[79].

Proof of lemma 3.11. Follows from lemma 3.9 and 3.10 as in Guaspari-Solovay[79].

This completes the proof of Theorem 3.3.

3.12. Remark.  $\Sigma_1$ -completeness is used only in the proof of Solovay's lemma, i.e; the the proof of lemma 2.2. It follows that

if, for some  $\Sigma_1$ -sound theory  $T \supseteq I\Delta_0$  we can get sentences  $L_i$   $i=0, \dots, n$  satisfying (1), ..., (5), we can embed finite  $R$ -models in  $T$ . This does not necessarily imply that we have arithmetical completeness for  $T$ , as  $R$  need not be arithmetically sound with respect to the interpretations in  $T$ .

### Bibliography:

- Guaspari, D. and R.M. Solovay, 'Rosser Sentences', *Annals of Mathematical Logic* 16, pp. 81-99, 1979.
- de Jongh, D.H.J., 'A Simplification of a Completeness Proof of Guaspari and Solovay', *Studia Logica* 46, pp. 187-192, 1987.
- Jumelet, M., On Solovay's Theorem, *Master's Thesis*, Department of Mathematics and Computer Science, University of Amsterdam, 1988.
- Smoryński, C., *Self-Reference and Modal Logic*, Springer, New York, 1985.
- Solovay, R.M., 'Provability interpretations of modal logic', *Israel Journal of Mathematics*, vol.25, pp. 287-304, 1976.
- Verbrugge, L.C., 'Does Solovay's Completeness Theorem extend to Bounded Arithmetic?', *Master's Thesis*, Department of Mathematics and Computer Science, University of Amsterdam, 1988.

## The ITLI Prepublication Series

1986

86-01  
86-02 Peter van Emde Boas  
86-03 Johan van Benthem  
86-04 Reinhard Muskens  
86-05 Kenneth A. Bowen, Dick de Jongh

86-06 Johan van Benthem

1987

87-01 Jeroen Groenendijk, Martin Stokhof  
87-02 Renate Bartsch  
87-03 Jan Willem Klop, Roel de Vrijer  
87-04 Johan van Benthem  
87-05 Víctor Sánchez Valencia  
87-06 Eleonore Oversteegen  
87-07 Johan van Benthem  
87-08 Renate Bartsch  
87-09 Herman Hendriks

1988

*Logic, Semantics and Philosophy of Language:*

LP-88-01 Michiel van Lambalgen  
LP-88-02 Yde Venema  
LP-88-03  
LP-88-04 Reinhard Muskens  
LP-88-05 Johan van Benthem  
LP-88-06 Johan van Benthem  
LP-88-07 Renate Bartsch  
LP-88-08 Jeroen Groenendijk, Martin Stokhof  
LP-88-09 Theo M.V. Janssen  
LP-88-10 Anneke Kleppe

*Mathematical Logic and Foundations:*

ML-88-01 Jaap van Oosten  
ML-88-02 M.D.G. Swaen  
  
ML-88-03 Dick de Jongh, Frank Veltman  
ML-88-04 A.S. Troelstra  
ML-88-05 A.S. Troelstra

*Computation and Complexity Theory:*

CT-88-01 Ming Li, Paul M.B. Vitanyi  
CT-88-02 Michiel H.M. Smid  
CT-88-03 Michiel H.M. Smid, Mark H. Overmars  
Leen Torenvliet, Peter van Emde Boas  
CT-88-04 Dick de Jongh, Lex Hendriks  
Gerard R. Renardel de Lavalette  
CT-88-05 Peter van Emde Boas  
CT-88-06 Michiel H.M. Smid  
  
CT-88-07 Johan van Benthem  
CT-88-08 Michiel H.M. Smid, Mark H. Overmars  
Leen Torenvliet, Peter van Emde Boas  
CT-88-09 Theo M.V. Janssen  
CT-88-10 Edith Spaan, Leen Torenvliet  
Peter van Emde Boas  
CT-88-11 Sieger van Denneheuvel  
Peter van Emde Boas

*Other prepublications:*

X-88-01 Marc Jumelet

1989

*Logic, Semantics and Philosophy of Language:*

LP-89-01 Johan van Benthem  
*Mathematical Logic and Foundations:*  
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

*Computation and Complexity Theory:*

CT-89-01 Michiel H.M. Smid  
CT-89-02 Peter van Emde Boas  
CT-89-03 Ming Li, Herman Neuféglise  
Leen Torenvliet, Peter van Emde Boas  
CT-89-04 Harry Buhrman, Leen Torenvliet  
CT-89-05 Pieter H. Hartel, Michiel H.M. Smid  
Leen Torenvliet, Willem G. Vree  
CT-89-06 H.W. Lenstra, Jr.

*Other prepublications:*

X-89-01 Marianne Kalsbeek  
X-89-02 G. Wagemakers  
X-89-03 A.S. Troelstra

The Institute of Language, Logic and Information  
A Semantical Model for Integration and Modularization of Rules  
Categorial Grammar and Lambda Calculus  
A Relational Formulation of the Theory of Types  
Some Complete Logics for Branched Time, Part I  
Well-founded Time, Forward looking Operators  
Logical Syntax

Type shifting Rules and the Semantics of Interrogatives  
Frame Representations and Discourse Representations  
Unique Normal Forms for Lambda Calculus with Surjective Pairing  
Polyadic quantifiers  
Traditional Logicians and de Morgan's Example  
Temporal Adverbials in the Two Track Theory of Time  
Categorial Grammar and Type Theory  
The Construction of Properties under Perspectives  
Type Change in Semantics:  
The Scope of Quantification and Coordination

Algorithmic Information Theory  
Expressiveness and Completeness of an Interval Tense Logic  
Year Report 1987  
Going partial in Montague Grammar  
Logical Constants across Varying Types  
Semantic Parallels in Natural Language and Computation  
Tenses, Aspects, and their Scopes in Discourse  
Context and Information in Dynamic Semantics  
A mathematical model for the CAT framework of Eurotra  
A Blissymbolics Translation Program

Lifschitz' Realizability  
The Arithmetical Fragment of Martin Löf's Type Theories with weak  $\Sigma$ -elimination  
Provability Logics for Relative Interpretability  
On the Early History of Intuitionistic Logic  
Remarks on Intuitionism and the Philosophy of Mathematics

Two Decades of Applied Kolmogorov Complexity  
General Lower Bounds for the Partitioning of Range Trees  
Maintaining Multiple Representations of  
Dynamic Data Structures  
Computations in Fragments of Intuitionistic Propositional Logic

Machine Models and Simulations (revised version)  
A Data Structure for the Union-find Problem  
having good Single-Operation Complexity  
Time, Logic and Computation  
Multiple Representations of Dynamic Data Structures  
Towards a Universal Parsing Algorithm for Functional Grammar  
Nondeterminism, Fairness and a Fundamental Analogy  
Towards implementing RL

On Solovay's Completeness Theorem

The Fine-Structure of Categorial Semantics

Explicit Fixed Points for Interpretability Logic  
Extending the Lambda Calculus with Surjective Pairing is conservative  
Rosser Orderings and Free Variables  
On the Proof of Solovay's Theorem

Dynamic Deferred Data Structures  
Machine Models and Simulations  
On Space efficient Solutions

A Comparison of Reductions on Nondeterministic Space  
A Parallel Functional Implementation of Range Queries

Finding isomorphisms between finite fields

An Orey Sentence for Predicative Arithmetic  
New Foundations. a Survey of Quine's Set Theory  
Index of the Heyting Nachlass



