Department of Philosophy - Utrecht University

Propositional combinations of Σ-sentences in Heyting's Arithmetic

A. Visser

Logic Group
Preprint Series
No. 117
July 1994

Utrecht Research
Institute for
Philosophy

ISBN 90-393-0713-x
ISSN 0929-0710
Dr. A. Visser, Editor

Propositional combinations of Σ-sentences in Heyting's Arithmetic

Explorations between intuitionistic propositional logic and intuitionistic arithmetic

Albert Visser

Version 1.0

ABSTRACT: This paper is mainly about Boolean combinations of Σ-formulas (B(Σ)-formulas) in HA. We prove a theorem guaranteeing that if a B(Σ)-formula is provable in HA, then a simpler one is also provable. Our theorem yields a characterization of the derived rules of HA for Σ-substitutions. Another application is a decision procedure for the closed fragment of the provability logic of HA. The proof of our theorem leads us to such varied subjects as NNIL-formulas, robust formulas and preservativity notions.

1 Introduction

1.1 The problem and its solution: This paper is about the simple question:

> *What more do we know, when we know that a Boolean (or, if you prefer, a Brouwerean) combination of Σ-sentences is provable in Heyting's Arithmetic?*

The answer takes the form:

> *We often know that a better Boolean combination of the same Σ-sentences is provable in Heyting's Arithmetic. Moreover the better Boolean combination can be found independently of the specific Σ-sentences under consideration.*

Here *better* has two components. It means 'stronger' in the sense that the stronger sentence implies the weaker one in the Intuitionistic Propositional Calculus (IPC). It also means 'simpler' in the sense that the simpler formula lies in a specific class NNIL, the class of propositional formulas with No Nestings of Implications to the

Left. The NNIL-formulas are modulo IPC-provable equivalence precisely the formulas preserved under sub-models in the usual Kripke semantics for IPC. The qualification *often* in the answer is just a gloss for:

unless our Boolean combination is already in NNIL.

We give a more formal statement of our result. Let A be a propositional formula. We let $\mathfrak{f}$ range over assignments of Σ-sentences to (at least) the propositional formulas occurring in A. We write '$A[\mathfrak{f}]$' for the result of substituting $\mathfrak{f}(p)$ for p in A. Our theorem says: there is an $A^* \in$ NNIL such that:

a) $IPC \vdash A^* \rightarrow A$, and

b) for all $\mathfrak{f}$: $HA \vdash A[\mathfrak{f}] \Rightarrow HA \vdash A^*[\mathfrak{f}]$.

A better answer to our question is not possible, at least if we abstract away from the specific Σ-sentences substituted. We will show:

c) For every propositional formula B such that $IPC \nvdash A^* \rightarrow B$, there is an $\mathfrak{f}$ such that:
$HA \vdash A^*[\mathfrak{f}]$ and $HA \nvdash B[\mathfrak{f}]$.

Or contraposed:

d) For every propositional formula B:
$\forall \mathfrak{f}\ (HA \vdash A^*[\mathfrak{f}] \Rightarrow HA \vdash B[\mathfrak{f}]) \Rightarrow IPC \vdash A^* \rightarrow B$.

From (a),(b) we get:

e) $\forall \mathfrak{f}\ (HA \vdash A[\mathfrak{f}] \Leftrightarrow HA \vdash A^*[\mathfrak{f}])$.

And hence (d) is equivalent with:

f) For every propositional formula B:
$\forall \mathfrak{f}\ (HA \vdash A[\mathfrak{f}] \Rightarrow HA \vdash B[\mathfrak{f}]) \Rightarrow IPC \vdash A^* \rightarrow B$.

So A^* is the strongest 'general' improvement of A. Of course, for specific $\mathfrak{f}$ often better improvements are possible, but A^* is the best one that works for all $\mathfrak{f}$. Note that (f) implies (a) by substituting A for B. So our results are summarized by (b) and (f).

1.2 What more is there in the paper? Our main result has two immediate applications. Firstly it yields a description of the derived rules of HA for Σ-substitutions. Secondly we can read off a decision procedure for the closed fragment of the provability logic of HA. Thus we solve a variant of Friedman's 35th problem (see Friedman[75]). (This solution occurs already in Visser[85])

Both for the smooth formulation of our main result and for its proof we have to develop some further material. Part of it involves the theory of certain classes of propositional formulas, like NNIL and ROB (to be introduced in the paper). A second part is the study of certain relations between formulas, called semi-consequence relations. The semi-consequence relations include derived rules and preservativity rela-

tions (dual to conservativity relations). One preservativity relation, namely Σ-preservativity (the 'dual' of Π-conservativity), will be studied in some detail in a long excursion (section 7).

The main technical result of the paper is the specification and verification of the NNIL-algorithm in section 5. Our results in both propositional logic and arithmetic are immediate consequences of the existence of this algorithm.

The contents of the paper is as follows:

Section 2	preliminaries to propositional logic
Section 3	consequence relations and preservativity notions
Section 4	robust formulas
Section 5	the NNIL-algorithm
Section 6	basic facts and notations for arithmetic
Section 7	digression: assorted facts on Σ-preservativity
Section 8	closure properties of Σ-preservativity over HA
Section 9	on Σ-substitutions
Section 10	the closed fragment of the provability logic of HA

1.3 Environment and history of the paper: This paper is part of the wider study of substitutions considered as a kind of semantics. We point to some examples of related work. First there is, of course, the project of provability and interpretability logic. See e.g. Smoryński[85], Boolos[93], Berarducci[90], Visser[90], EA[91], Shavrukov[93], Zambella[94]. Then there is the work on derived rules by Rybakov. See e.g. Rybakov[92]. For a systematical exploration of the notion of derived rule, see FHY[92]. Finally there is the most immediate environment: research on substitutions in intuitionistic logic of propositional formulas or arithmetical sentences. See e.g. De Jongh[82], Smoryński[73], Leivant[75,80,81], Van Oosten[91], DV[94], DC[?].

The present paper is, partly, a remake of the unpublished paper *Evaluation, provably deductive equivalence in Heyting's Arithmetic* (Visser[85]). The work in Visser[85] was inspired by Dick de Jongh's results on propositional formulas of one variable, reported in De Jongh[82]. Since Visser[85] was written, the development of interpretability logic took place (see e.g. Berarducci[90] and Visser[90]). Interpretability logic inspired me to give the notion of preservativity a more prominent role in the presentation.

In Visser[85] we proved the identity of certain formula classes NNIL and ROB. Johan van Benthem gave, independently, around 1984 an alternative proof of the fact that

NNIL=ROB (see his Van Benthem[91]). Gerard Renardel proved (also around 1984, see his Renardel[86]), that NNIL satisfies both left and right interpolation. The work of both van Benthem and Renardel is contained and extended in DRVV[94]. This last paper should be considered as a companion paper of the present paper.

Dick de Jongh and Albert Visser recently reopened the research on Heyting algebras of arithmetical theories. A number of concerns of the present paper (derived rules, exact formulas) reappear in their paper DV[94], which is another companion paper of the present paper.

1.4 Acknowledgements: In various stages of research I benefited from the work, the wisdom and/or the advice of: Johan van Benthem, Dirk van Dalen, Dick de Jongh, Karst Koymans, Piet Rodenburg, Volodya Shavrukov, Rick Statman, Anne Troelstra and Domenico Zambella.

1.5 Prerequisites: Some knowledge of Troelstra[73] or TV[88a,b] is certainly beneficial. At some places I will make use of results from Visser[82] and DV[94].

2 Preliminaries to propositional logic: $\mathfrak{P},\mathfrak{Q},\ldots$ will be *sets of propositional variables*. **p,q,r,...** will be *finite sets of propositional variables*. We define $\mathfrak{L}(\mathfrak{P})$ as the smallest set $\mathfrak{S}$ such that:

- $\mathfrak{P}\subseteq\mathfrak{S}$, $\top,\bot\in\mathfrak{S}$,
- If $A,B\in\mathfrak{S}$, then $(A\wedge B)$, $(A\vee B)$, $(A\rightarrow B)\in\mathfrak{S}$.

SUB(A) is the set of subformulas of A. By convention we will count $\bot$ a subformula of any A. PV(A) is the set of propositional variables occurring in A.

We suppose that the reader is familiar with Kripke models for IPC (see TV[88a], or Smoryński [73]). Our treatment here is mainly to fix notations. A model is a structure $\mathbb{K} = \langle K,\leq,\Vdash,\mathfrak{P}\rangle$, where K is a non-empty set of nodes, $\leq$ is a partial ordering. $\Vdash$ is the atomic forcing relation for $\mathfrak{P}$: it is a relation between nodes and propositional atoms in $\mathfrak{P}$, satisfying: $k\leq k'$ and $k\Vdash p \Rightarrow k'\Vdash p$ (persistence). $\mathbb{K}$ is a *$\mathfrak{P}$-model* if $\mathfrak{P}_{\mathbb{K}}=\mathfrak{P}$. Mod($\mathfrak{P}$) is the set of $\mathfrak{P}$-models.

Consider $\mathbb{K} = \langle K,\leq,\Vdash,\mathfrak{P}\rangle$. We define $k\Vdash A$ for $A\in\mathfrak{L}(\mathfrak{P})$ in the standard way. We write $\mathbb{K}\Vdash A$ for: $\forall k\in K\ k\Vdash A$.

A model $\mathbb{K}$ is *finite* if both $K_{\mathbb{K}}$ and $\mathfrak{P}_{\mathbb{K}}$ are finite.

A *rooted model* $\mathbb{K}$ is a structure $\langle K,\mathfrak{b},\leq,\vDash,\mathfrak{P}\rangle$, where $\langle K,\leq,\vDash,\mathfrak{P}\rangle$ is a model and where $\mathfrak{b}\in K$ is the bottom element w.r.t. $\leq$.. The set of rooted models is Rmod.

For any $k\in K$ $\mathbb{K}[k]$ is the model $\langle K',k,\leq',\vDash',\mathfrak{P}\rangle$, where $K':= \uparrow k := \{k'|k\leq k'\}$ and where $\leq'$ and $\vDash'$ are the restrictions of $\leq$ respectively $\vDash$ to K'. (We will often simply write $\leq$ and $\vDash$ for $\leq'$ and $\vDash'$.)

2.1 The Henkin construction: A set $X\subseteq\mathfrak{L}(\mathfrak{P})$ is *adequate* if it is closed under subformulas and contains $\bot$. A set Γ is X-saturated if:

(i) $\Gamma\subseteq X$, (ii) $\Gamma\nvdash\bot$, (iii) $\Gamma\vdash A$, $A\in X\Rightarrow A\in\Gamma$,

(iv) $\Gamma\vdash(B\vee C)$, $(B\vee C)\in X\Rightarrow B\in\Gamma$ or $C\in\Gamma$.

The Henkin model $\mathbb{H}_X(\mathfrak{P})$ is the $\mathfrak{P}$-model with as nodes the X-saturated sets Δ and as accessibily relation $\subseteq$. The atomic forcing in the nodes is given by: $\Gamma\vDash p\Leftrightarrow p\in\Gamma$. We have by a standard argument: for $A\in X$: $\Gamma\vDash A\Leftrightarrow A\in\Gamma$. Note that if X is finite, then $\mathbb{H}_X(\mathfrak{P})$ is finite. A direct consequence of the Henkin construction is the Kripke Completeness Theorem. Let $\mathfrak{P}\supseteq PV(A)$, then:

IPC$\vdash A\Leftrightarrow$ for all (finite) $\mathfrak{P}$-models $\mathbb{K}$: $\mathbb{K}\vDash A$.

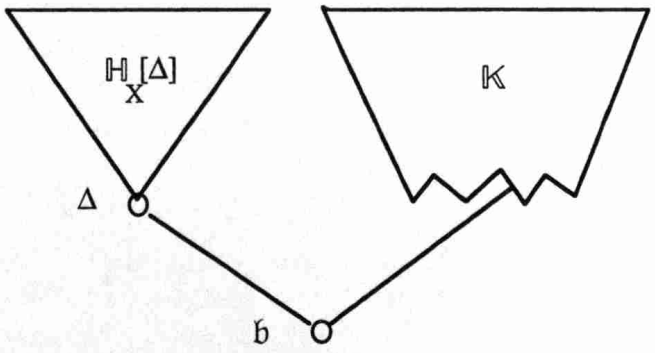
2.2 Futher Definitions

i) Let $\mathbf{K}$ be a set of $\mathfrak{P}$-models. M($\mathbf{K}$) is the $\mathfrak{P}$-model with nodes $\langle k,\mathbb{K}\rangle$ for $k\in K_{\mathbb{K}}$, $\mathbb{K}\in\mathbf{K}$ and ordering: $\langle k,\mathbb{K}\rangle\leq\langle m,\mathbb{M}\rangle :\Leftrightarrow \mathbb{K}=\mathbb{M}$ and $k\leq_{\mathbb{K}} m$. As atomic forcing we define: $\langle k,\mathbb{K}\rangle\vDash p :\Leftrightarrow k\vDash_{\mathbb{K}} p$. (In practice we will forget the second components of the new nodes, pretending the domains to be already disjoint.)

ii) Let $\mathbb{K}$ be a $\mathfrak{P}$-model. B($\mathbb{K}$) is the rooted $\mathfrak{P}$-model obtained by adding a new bottom $\mathfrak{b}$ to $\mathbb{K}$ and by taking: $\mathfrak{b}\vDash p :\Leftrightarrow \mathbb{K}\vDash p$. We put Glue($\mathbf{K}$) := B(M($\mathbf{K}$)). ○

We will assume below that $\mathfrak{P}$ is fixed. We will often notationally suppress it.

2.3 Push Down Lemma: Let X be adequate. Suppose Δ is X-saturated and $\mathbb{K}\vDash\Delta$. Then Glue($\mathbb{H}_X[\Delta],\mathbb{K}$)$\vDash\Delta$.



Proof: We show by induction on $A \in X$ that $\mathfrak{b} \Vdash A \Leftrightarrow A \in \Delta$. The cases of atoms, conjunction and disjunction are trivial. If $(B \to C) \in X$ and $\mathfrak{b} \Vdash (B \to C)$, then $\Delta \Vdash (B \to C)$ and hence $(B \to C) \in \Delta$. Conversely suppose $(B \to C) \in \Delta$. If $\mathfrak{b} \nVdash B$, we are easily done. If $\mathfrak{b} \Vdash B$, then, by the Induction Hypothesis: $B \in \Delta$, hence $C \in \Delta$ and, by the Induction Hypothesis: $\mathfrak{b} \Vdash C$. ❑

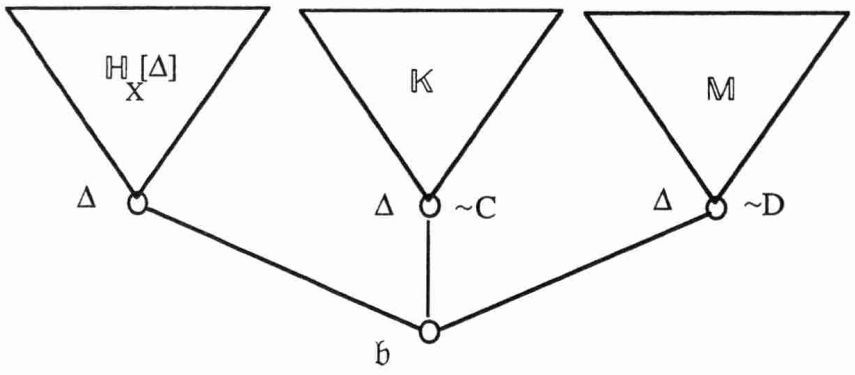
We say that Δ is ($\mathfrak{P}$-)*prime* if it is consistent and:

for every $(C \vee D) \in \mathfrak{L}(\mathfrak{P})$: $\Delta \vdash (C \vee D) \Rightarrow \Delta \vdash C$ or $\Delta \vdash D$.

A formula A is prime if $\{A\}$ is prime.

2.4 **Theorem:** Suppose X is adequate and Δ is X-saturated. Then Δ is prime.

Proof: Δ is consistent by definition. Suppose $\Delta \vdash C \vee D$ and $\Delta \nvdash C$ and $\Delta \nvdash D$. Suppose $\mathbb{K} \Vdash \Delta$, $\mathbb{K} \nVdash C$, $\mathbb{M} \Vdash \Delta$ and $\mathbb{M} \nVdash D$. Consider Glue($\mathbb{H}_X(\Delta), \mathbb{K}, \mathbb{M}$). By 2.3 we have: $\mathfrak{b} \Vdash \Delta$. On the other hand, by persistence: $\mathfrak{b} \nVdash C$ and $\mathfrak{b} \nVdash D$. Contradiction. ❑



(~C exhibited next to a node means that *C is not forced*; this is not to be confused with ¬C exhibited next to a node, which means that *¬C is forced*.)

2.5 **Theorem:** Consider any formula A. The formula A can be written (modulo IPC-provable equivalence) as a disjunction of prime formulas C. Moreover these C are conjunctions of implications and propositional variables in SUB(A).

Proof: Consider a SUB(A)-saturated Δ. Let IP(Δ) be the set of implications and atoms of Δ. It is easily seen that IPC $\vdash \bigwedge \mathrm{IP}(\Delta) \leftrightarrow \bigwedge \Delta$. Take:

$D := \bigvee \{ \bigwedge \mathrm{IP}(\Delta) \mid \Delta \text{ is SUB(A)-saturated and } A \in \Delta \}$.

Trivially: IPL $\vdash D \to A$. On the other hand if IPC $\nvdash A \to D$, then by a standard construction there is a SUB(A)-saturated set Γ such that $A \in \Gamma$ and $\Gamma \nVdash D$. Quod non. ❑

2.5.1 Remark: Note that in the definition of D in 2.5, we can restrict ourselves to the $\subseteq$-minimal sub(A)-saturated Δ with $A \in \Delta$.

We define a measure of complexity ρ, which counts the left-nesting of $\rightarrow$, as follows:

- $\rho(p) := \rho(\bot) := \rho(\top) := 0$
- $\rho(A \wedge B) := \rho(A \vee B) := \max(\rho(A), \rho(B))$
- $\rho(A \rightarrow B) := \max(\rho(A)+1, \rho(B))$.

$NNIL(\mathfrak{P}) := \{A \in \mathfrak{L}(\mathfrak{P}) \mid \rho(A) \leq 1\}$. In other words NNIL is the class of formulas without nestings of implications to the left. An example of as NNIL-formula is:

$$(p \rightarrow (q \vee (s \rightarrow t))) \wedge ((q \vee r) \rightarrow s)).$$

It is easy to see that modulo IPC-provable equivalence each NNIL-formula can be rewritten to a $NNIL_0$-formula, i.e. a formula in which in front of implications only single atoms occur. For more information about NNIL, see DV[94] and DRVV[94].

3 Consequence relations & preservativity notions: An important tool of the present paper is the use of *semi-consequence relations* (defined below) and *preservativity relations* (defined below). '$\rhd$' will range over semi-consequence relations. Let $\vdash$ stand for derivability in IPC.

Let $\mathfrak{B}$ be a language (for propositional or predicate logic) and let T be a theory in $\mathfrak{B}$. A *semi-consequence relation on $\mathfrak{B}$ over T* is a binary relation on $\mathfrak{B}$ satisfying:

A1 $A \vdash_T B \Rightarrow A \rhd B$

A2 $A \rhd B$ and $B \rhd C \Rightarrow A \rhd C$

A3 $C \rhd A$ and $C \rhd B \Rightarrow C \rhd (A \wedge B)$.

The name 'semi-consequence relation' is ad hoc in this paper. We take:

- $A \equiv B :\Leftrightarrow A \rhd B$ and $B \rhd A$.

If we don't specify the theory with the semi-consequence relation, it's always supposed to be over IPC.

A further salient principle is:

B1 $A \rhd C$ and $B \rhd C \Rightarrow (A \vee B) \rhd C$.

A relation satisfying A1-A3,B1 is called a *nearly-consequence* relation. Note that $\vdash_T$ is a nearly-consequence relation over T.

3.1 Conventions: We will use $X \vdash Y$, $X \rhd Y$ for respecively $\wedge X \vdash \vee Y$ and $\wedge X \rhd \vee Y$, where X and Y are finite sets of formulas. ($\wedge \emptyset := \top$, $\vee \emptyset := \bot$.). We treat implications similarly.

We write X,Y for: $X \cup Y$, X,A for $X \cup \{A\}$, etc. ○

Nearly-consequence relations over T can be alternatively described in Genzen style as follows. Nearly-consequence relations are consequence relations satifying:

A1' $X \vdash_T Y \Rightarrow X \rhd Y$

Thin $X \rhd Y \Rightarrow X,Z \rhd Y,U$

Cut $X \rhd Y,A$ and $Z,A \rhd U \Rightarrow X,Z \rhd Y,U$

We take the permutation rules to be implicit in the set notation. We leave it to the reader to check the equivalence of the Genzen style principles with A1-A3,B1.

3.2 **Fact:** Let $\rhd$ be a semi-consequence relation on $\mathfrak{L}(\mathfrak{P})$. Suppose there is a $\Phi:\mathfrak{L}(\mathfrak{P}) \to \mathfrak{L}(\mathfrak{P})$, such that for all $A,B \in \mathfrak{L}(\mathfrak{P})$: $A \rhd B \Leftrightarrow \Phi(A) \vdash B$. Then:

i) If for all $B \in \mathfrak{L}(\mathfrak{P})$: $A \rhd B \Leftrightarrow C \vdash B$, then $\vdash C \leftrightarrow \Phi(A)$.

ii) Φ considered as a function on the Heyting algebra of IPC is a co-closure operation, i.e.:

- $\Phi(A) \vdash A$ (co-inductivity)
- $A \vdash B \Rightarrow \Phi(A) \vdash \Phi(B)$ (monotonicity)
- $\vdash \Phi(\Phi(A)) \leftrightarrow \Phi(A)$ (idempotency).

In fact we have more than monotonicity, viz:

- $A \rhd B \Leftrightarrow \Phi(A) \vdash \Phi(B)$.

iii) Suppose $\rhd$ is a nearly-consequence relation. Then $\vdash \Phi(A \vee B) \leftrightarrow (\Phi(A) \vee \Phi(B))$.

Proof: The simple proofs are left to the reader. ❑

Some principles involving implication play an important role in the paper. To introduce them we need a syntactical operation. We define the operation [.](.) on propositional formulas as follows:

- $[B]p := p$, $[B]\top := \top$, $[B]\bot := \bot$,
- [.](.) commutes with $\wedge$ and $\vee$,
- $[B](C \to D) := (B \to (C \to D))$.

Note that [.](.) does not preserve provable equivalence in the second component. Note also that: $\vdash B \to ([B]C \leftrightarrow C)$. Define: $[B]X := \{[B]C \mid C \in X\}$.

We have the following principles for implication for semi-consequence relations on $\mathfrak{L}(\mathfrak{P})$:

B2 $A \rhd B \Rightarrow (p \to A) \rhd (p \to B)$

B3 Let X be a finite set of implications and let $Y := \{C \mid (C \to D) \in X\} \cup \{B\}$. Take $A := \bigwedge X$ then: $(A \to B) \rhd [A]Y$

We give an example of an instance of B3. Let e.g.

$A := ((p \to q) \wedge (r \to (s \vee t)))$, $C := (A \to (u \vee (p \to r)))$,

then we have:

$C \rhd ([A]p \vee [A]r \vee [A](u \vee (p \to r))$, i.e.:

$C \rhd (p \vee r \vee u \vee (A \to (p \to r))$.

3.3 **Open problem:** Is it possible to axiomatize B3 by a finite number of traditional schemes? We do not allow finite sets or syntactical operations to occur in such schemes. However we *do* allow schematic letters ranging over arbitrary formulas and schematic letters ranging over propositional variables. (So e.g. B2 is acceptable as a scheme.) ○

We say that a relation satisfying A1-A3,B1-B3 is a *σ-relation.*

Consider again any language $\mathfrak{B}$ of propositional or predicate logic. Let $X \subseteq \mathfrak{B}$ and let T be any theory in that language. Let $\mathfrak{F}$ be a set of functions from $\mathfrak{P}$ to $\mathfrak{B}$. We write:

- For $A,B \in \mathfrak{B}$: $A \rhd_{T,X} B :\Leftrightarrow \forall C \in X\, (C \vdash_T A \Rightarrow C \vdash_T B)$,
- Let $A \in \mathfrak{L}(\mathfrak{P})$. $A[f]$ is the result of substituting $f(p)$ for p in A for each $p \in \mathfrak{P}$.
- For $A,B \in \mathfrak{L}(\mathfrak{P})$: $A \rhd_{T,X,\mathfrak{F}} B :\Leftrightarrow \forall f \in \mathfrak{F}\, A[f] \rhd_{T,X} B[f]$.

If $Y \subseteq \mathfrak{B}$ and $\mathfrak{F} = Y^{\mathfrak{P}}$, we will write $\rhd_{T,X,Y} B$ for $\rhd_{T,X,\mathfrak{F}}$. If X or $\mathfrak{F}$ are singletons we will omitt the singleton brackets. If $\mathfrak{B} = \mathfrak{L}(\mathfrak{P})$ and T=IPC, we will often omitt 'T' in the index. We will call $\rhd_{T,X,\mathfrak{F}}$ *T,X,$\mathfrak{F}$-preservativity*, etcetera. We will call the $\rhd_{T,X,\mathfrak{F}}$ and the $\rhd_{T,X}$ *preservativity relations.* We will call the $\rhd_{T,X}$ *pure preservativity relations.*

Clearly $\rhd_{T,X}$ is a semi-consequence relation over T and $\rhd_{T,X,\mathfrak{F}}$ is a semi-consequence relation over IPC.

Below we will provide a number of motivating examples for our definitions.

3.4 **Remark:** It is instructive to compare preservativity with conservativity. Define:

- For $A,B \in \mathfrak{B}$: $A \rhd^*_{T,X} B :\Leftrightarrow \forall C \in X\, (B \vdash_T C \Rightarrow A \vdash_T C)$,

So $A \rhd^*_{T,X} B$ means that T+B is conservative over T+A w.r.t. X. For *classical* theories T we have:

☆ $A \rhd_{T,X} B \Leftrightarrow \neg B \rhd^*_{T,\neg X} \neg A$, where $\neg X := \{\neg C | C \in X\}$.

Thus classically preservativity is a superfluous notion. Constructively, however, the reduction given in ☆ does not work.

Note that conservativity as a relation between sentences over a theory is a natural extension of the notion of conservativity between theories. There is no analogue of this for preservativity. ○

A formula A of $\mathfrak{B}$ is T-*prime* if for any finite set of $\mathfrak{B}$-formulas Z:

$A\vdash_T Z \Rightarrow \exists B\in Z\ A\vdash_T B$.

Note that if A is T-prime, then $A\nvdash_T \bot$. A class of formulas X is *weakly T-disjunctive* if every $A\in X$ is equivalent to the disjunction of a finite set of T-prime formulas Y, with $Y\subseteq X$.

For any class of formulas X, let DISJ(X) be the closure of X under arbitrary conjunctions.

In the next fact we collect a number of noteworthy small facts on preservativity.

3.5 Fact

i) $\rhd_{T,X}$ is a semi-consequence relation over T. $\rhd_{T,X,\mathfrak{F}}$ is a semi-consequence relation over IPC.

ii) Suppose X is weakly T-disjunctive, then $\rhd_{T,X}$ and $\rhd_{T,X,\mathfrak{F}}$ are nearly-consequence relations.

iii) Suppose X is closed under conjunction, then:

$C\in X$ and $A\rhd_{T,X}B \Rightarrow (C\to A)\rhd_{T,X}(C\to B)$.

iv) Let range($\mathfrak{F}$) be the union of the ranges of the elements of $\mathfrak{F}$. Suppose range($\mathfrak{F}$)$\subseteq X$ and X is closed under conjunction, then $\rhd_{T,X,\mathfrak{F}}$ satisfies B2.

v) Suppose $A\in X$, then: $A\rhd_{T,X}B \Leftrightarrow A\vdash_T B$.

vi) Suppose $X\subseteq Y$ and $\mathfrak{F}\subseteq\mathfrak{G}$, then $\rhd_{T,Y}\subseteq\rhd_{T,X}$ and $\rhd_{T,Y,\mathfrak{G}}\subseteq\rhd_{T,X,\mathfrak{F}}$.

vii) $\rhd_{T,DISJ(X)}=\rhd_{T,X}$ and hence $\rhd_{T,DISJ(X),\mathfrak{F}}=\rhd_{T,X,\mathfrak{F}}$.

viii) Let $\mathfrak{id}:\mathfrak{P}\to\mathfrak{L}(\mathfrak{P})$ be the function with $\mathfrak{id}(p)=p$. Then: $\rhd_{IPC,X,\mathfrak{id}}=\rhd_{IPC,X}$.

Proof: We treat (ii) and (iii). (ii) Suppose X is weakly T-disjunctive and $A\rhd_{T,X}C$ and $B\rhd_{T,X}C$. Let E be any element of X. Suppose Y is a finite set of T-prime formulas from X such that E is equivalent to the disjunction of Y. We have:

$$\begin{aligned} E\vdash_T A\vee B &\Rightarrow \bigvee Y\vdash_T A\vee B\\ &\Rightarrow \forall F\in Y\ F\vdash_T A\vee B\\ &\Rightarrow \forall F\in Y\ F\vdash_T A \text{ or } F\vdash_T B\\ &\Rightarrow \forall F\in Y\ F\vdash_T C\\ &\Rightarrow \bigvee Y\vdash_T C\\ &\Rightarrow E\vdash_T C \end{aligned}$$

B1 for $\rhd_{T,X,\mathfrak{F}}$ is immediate from B1 for $\rhd_{T,X}$.

iii) Suppose X is closed under conjunction. Suppose $A \rhd B$. Let $C,E \in X$ and suppose $E \vdash (C \to A)$. Then $(E \wedge C) \vdash A$ and hence $(E \wedge C) \vdash B$. Ergo $E \vdash (C \to B)$. ❑

3.6 Example 1: Take $X := \{ \top \}$. We have:

For $A,B \in \mathfrak{B}$: $A \rhd_{T,\top} B \Leftrightarrow (\vdash_T A \Rightarrow \vdash_T B)$,

For $A,B \in \mathfrak{L}(\mathfrak{P})$: $A \rhd_{T,\top,\mathfrak{F}} B \Leftrightarrow \forall f \in \mathfrak{F}\ (\vdash_T A[f] \Rightarrow \vdash_T B[f])$.

Thus $\rhd_{T,\top}$ is the relation of *deductive consequence* for T and $\rhd_{T,\top,\mathfrak{F}}$ is the relation of being a (propositional) *derived rule* for T w.r.t. $\mathfrak{F}$. We mention the important result due to Rybakov (see Rybakov[92]) that $\rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})}$ is decidable.

Note that if T has the disjunction property, then $\rhd_{T,\top}$ and $\rhd_{T,\top,\mathfrak{F}}$ satisfy B1.

We repeat a well known observation.

3.6.1 Theorem: Suppose for all $C \in \mathfrak{B}$: $\vdash C \Leftrightarrow \forall f \in \mathfrak{F} \vdash_T C[f]$. Moreover suppose $\mathfrak{G} \subseteq \mathfrak{L}(\mathfrak{P})^{\mathfrak{P}}$ and for any $g \in \mathfrak{G}$ and $f \in \mathfrak{F}$: $g \circ f \in \mathfrak{F}$. (We read composition in the order of application, so that $A[g][f] = A[g \circ f]$.) Then: $\rhd_{T,\top,\mathfrak{F}} \subseteq \rhd_{IPC,\top,\mathfrak{G}}$.

Proof: Suppose $A \rhd_{T,\top,\mathfrak{F}} B$ and $\vdash A[g]$. Consider any $f \in \mathfrak{F}$. Clearly $\vdash_T A[g][f]$, i.e. $\vdash_T A[g \circ f]$. Since $g \circ f \in \mathfrak{F}$, we find: $\vdash_T B[g \circ f]$. Hence for all $f \in \mathfrak{F}$: $\vdash_T B[g][f]$ and hence: $\vdash B[g]$. ❑

Consider e.g. HA. Let $\mathfrak{A}$ be the language of HA, let Σ be the set of Σ-sentences and let $B(\Sigma)$ be the set of Boolean (Brouwerean) combinations of Σ-sentences. We have:

$\divideontimes$ for all $C \in \mathfrak{A}$: $\vdash C \Leftrightarrow \forall f \in \Sigma^{\mathfrak{P}} \vdash_{HA} C[f]$.

$\divideontimes$ is De Jongh's completeness theorem for IPC w.r.t. Σ-substitutions in HA. There are many different proofs of $\divideontimes$, see e.g. Smoryński [73] or Visser[85] or DV[94]. Let $\mathfrak{F} := B(\Sigma)^{\mathfrak{P}}$ and $\mathfrak{G} := \mathfrak{L}(\mathfrak{P})^{\mathfrak{P}}$ and T:=HA in 3.6.1. We find: $\rhd_{HA,\top,B(\Sigma)} \subseteq \rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})}$. In DV[94] (theorem 6.2) it is shown that $\rhd_{HA,\top,B(\Sigma)} = \rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})}$. So we have:

$\dotplus$ $\rhd_{HA,\top,\mathfrak{A}} \subseteq \rhd_{HA,\top,B(\Sigma)} = \rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})} \subseteq \rhd_{HA,\top,\Sigma}$.

The main result of the present paper implies that $(\neg\neg p \to p) \rhd_{HA,\top,\Sigma} (p \vee \neg p)$. On the other hand we have: $IPC \vdash (\neg\neg p \to p)[p := \neg q]$, but not $IPC \vdash (p \vee \neg p)[p := \neg q]$. So the second inclusion of $\dotplus$ is strict. It is open whether the first inclusion of $\dotplus$ is strict.

In 3.8 we will see that $\rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})}$ is extensionally equal to a pure preservativity relation. We will see in 9.2, that $\rhd_{HA,\top,\Sigma}$ can be alternatively described as a pure preservativity relation.

3.6.2 Reformulation of our main result: The result we are aiming at in this paper is (a)+(e) of the introduction. In our new notation this becomes:

there is an $A^* \in NNIL$ such that:

b) $A \rhd_{HA, \top, \Sigma} A^*$.

f) for all $B \in \mathfrak{L}(\mathfrak{P})$: $A \rhd_{HA, \top, \Sigma} B \Rightarrow \vdash A^* \rightarrow B$.

We show that (b)+(f) is equivalent to:

g) for all $B \in \mathfrak{L}(\mathfrak{P})$: $A \rhd_{HA, \top, \Sigma} B \Leftrightarrow \vdash A^* \rightarrow B$.

Thus our main result takes the form of a characterization of $\rhd_{HA, \top, \Sigma}$.

Proof: "(b)+(f)⇒(g)" Suppose (b) and (f). We prove (g). From left to right is trivial. Suppose $\vdash A^* \rightarrow B$. Then by A1: $A^* \rhd_{HA, \top, \Sigma} B$. Hence by (b) and A2: $A \rhd_{HA, \top, \Sigma} B$.

"(g)⇒(b)+(f)" is trivial. ❑((b)+(f)⇔(g))

By 3.2(i) and (ii) we may conclude that (.)* gives unique values modulo IPC-provable equivalence and that (.)* is a co-closure operation on the Heyting algebra of IPC. We will see later that $\rhd_{HA, \top, \Sigma}$ satisfies B1 and that, hence, (.)* commutes with disjunction according to 3.2(iii) ○

3.7 Example 2: Take $X := \mathfrak{B}$, $\rhd_{T,X}$ is *T-provable consequence* or $\vdash_T$.

3.8 Example 3: We show that the notion of *derived rule for IPC* coincides with a pure preservativity relation.

An $\mathfrak{L}(\mathbf{p})$-formula A is *IPC,p-exact* if there is an $f \in \mathfrak{L}(\mathfrak{P})^{\mathbf{p}}$ such that:

☆ for all $B \in \mathfrak{L}(\mathbf{p})$: $\vdash B[f] \Leftrightarrow A \vdash B$.

Let's say that the class of IPC,**p**-exact formulas is EX(**p**). Let EX be the union of the EX(**p**)'s for all finite subsets **p** of $\mathfrak{P}$.

De Jongh & Visser show (see: DV[94], Theorem 2.3) that for any $f \in \mathfrak{L}(\mathfrak{P})^{\mathbf{p}}$ we can find an $A := A_f \in \mathfrak{L}(\mathbf{p})$ satisfying ☆. It is easy to see that A_f is unique modulo IPC-provable equivalence. Let **q** be the set of variables occurring in the range of f. Let f^+ be the result of putting $f^+(p) := f(p)$ if $p \in \mathbf{p}$, $f^+(p) := p$ otherwise. Inspection of the argument by de Jongh and Visser reveals that A satisfies the stronger:

✪ for all $B \in \mathfrak{L}(\mathbf{p} \cup (\mathfrak{P}/\mathbf{q}))$: $\vdash B[f^+] \Leftrightarrow A \vdash B$.

To see that the variable condition is needed, let e.g. $\mathbf{p} := \{p\}$, $f := [p := \neg q]$. It is easy to see that $A_f = (\neg\neg p \rightarrow p)$. Let $B := (\neg q \rightarrow p)$. $B[f^+] = (\neg q \rightarrow \neg q)$ and so $\vdash B[f^+]$. On the other hand $(\neg\neg p \rightarrow p) \nvdash (\neg q \rightarrow p)$. (The proof in DV[94], employs Pitts's Uniform Interpolati-

on Theorem for IPC (see Pitts[92]). Inspecting the proof it can be seen that the variable condition is analogous to the one of $\exists$-elimination.)

We show: $\rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})} = \rhd_{IPC,EX}$.

Suppose $B \rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})} C$ and $E \in EX$, say $E \in EX(\mathbf{p})$ and let $f \in \mathfrak{L}(\mathfrak{P})^{\mathbf{p}}$ witness the fact that $E \in EX(\mathbf{p})$. As is easily seen (by permuting $\mathfrak{P}$) we can arrange that the variables in the elements of the range of f are disjoint from $PV(B) \cup PV(C)$. Construct f^+ as above. We find using ❂:

$$E \vdash B \Rightarrow \vdash B[f^+] \Rightarrow \vdash C[f^+] \Rightarrow E \vdash C.$$

Conversely suppose $B \rhd_{IPC,EX} C$ and consider $f \in \mathfrak{L}(\mathfrak{P})^{\mathfrak{P}}$. Take $\mathbf{p} := PV(B) \cup PV(C)$ and $g := f|\mathbf{p}$ and $A := A_g$. Then:

$$\vdash B[f] \Rightarrow \vdash B[g] \Rightarrow A \vdash B \Rightarrow A \vdash C \Rightarrow \vdash C[g] \Rightarrow \vdash C[f].$$

In the cases that $\mathfrak{P}=\emptyset$ and $\mathfrak{P}=\{p\}$, $\rhd_{IPC,EX}$ is completely understood. The case that $\mathfrak{P}=\{p\}$ is the subject of De Jongh[82]. He shows that the exact formulas in one variable are precisely: p, $\neg p$, $\neg\neg p$, $\neg\neg p \rightarrow p$, $\top$. It follows that DISJ(EX) is finite (mod IPC). If we set $\Phi(A)$:= the disjunction of all EX-formulas that IPC-imply A, we get: $A \rhd_{IPC,EX} B \Leftrightarrow \Phi(A) \vdash B$. ❍

3.9 Remark: It is not dificult to see that every NNIL($\mathbf{p}$)-formula can be written as a finite disjunction of prime NNIL($\mathbf{p}$)-formulas. In DV[94] (Theorem 2.8) it is shown that prime NNIL($\mathbf{p}$)-formulas are IPC,$\mathbf{p}$-exact. So modulo IPC-provable equivalence: $NNIL(\mathbf{p}) \subseteq DISJ(EX(\mathbf{p}))$. Ergo by 3.5(vii) and 3.8:

$$\rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})} = \rhd_{IPC,EX} = \rhd_{IPC,DISJ(EX)} \subseteq \rhd_{IPC,NNIL}.$$

We leave it to the reader to show that the last inclusion is proper. (We will show in 5.2 and 9.2 that $\rhd_{IPC,NNIL} = \rhd_{HA,\top,\Sigma}$. So our result here coheres with ✣ of 3.6.) ❍

3.10 Example 4: Suppose $\mathfrak{Q} \subseteq \mathfrak{P}$. Consider $\rhd_{IPC,\mathfrak{L}(\mathfrak{Q})}$. By 2.5, $\mathfrak{L}(\mathfrak{Q})$ is weakly IPC-disjunctive. Hence, by 3.5(ii), $\rhd_{IPC,\mathfrak{L}(\mathfrak{Q})}$ is a nearly consequence relation. Note that $\rhd_{IPC,\mathfrak{L}(\mathfrak{Q})}$ also satifies:

if $C \in \mathfrak{L}(\mathfrak{Q})$: $A \rhd B \Rightarrow (C \rightarrow A) \rhd (C \rightarrow B)$,

if $p \notin \mathfrak{L}(\mathfrak{Q})$: $(p \rightarrow A) \rhd A[p := \top]$.

Andrew Pitts, in his Pitts[92], shows that for every $A \in \mathfrak{L}(\mathfrak{P})$, there is a formula $(\forall A) \in \mathfrak{L}(\mathfrak{Q})$, such that for all $B \in \mathfrak{L}(\mathfrak{P})$: $A \rhd_{IPC,\mathfrak{L}(\mathfrak{Q})} B \Leftrightarrow \forall A \vdash B$. We can read $\forall A$ informally as the result of universally quantifying out the propositional variables not in $\mathfrak{Q}$. Note that we may apply 3.2 to $\forall$.

3.11 List of results on preservativity: We close this section by listing a number of relationships between various preservativity relations.

1) $\rhd_{IPC,\mathfrak{L}(\mathfrak{P})}$ is the minimal preservativity relation and:
$\rhd_{IPC,\mathfrak{L}(\mathfrak{P})} \subset^{(\alpha)} \rhd_{HA,\top,\mathfrak{A}} \subseteq^{(\beta)} \rhd_{IPC,EX} \subset^{(\gamma)} \rhd_{IPC,NNIL} \subset \rhd_{IPC,\top}$.
   - α) Non-identity follows from the presence of non-trivial derived rules for HA, like IP, i.e. $(\neg p \rightarrow (q \vee r)) \rhd_{HA,\top,\mathfrak{A}} ((\neg p \rightarrow q) \vee (\neg p \rightarrow r))$.
   - β) Immediate from (3) below.
   - γ) Non-identity follows by: $(\neg\neg p \rightarrow p) \rhd_{IPC,NNIL} (p \vee \neg p)$ and the fact that $(\neg\neg p \rightarrow p)$ is IPC,p-exact by the substitution $[p:=\neg q]$. Hence we do not have $(\neg\neg p \rightarrow p) \rhd_{IPC,EX} (p \vee \neg p)$.
2) $\rhd_{IPC,\mathfrak{L}(\mathfrak{P}),\mathfrak{L}(\mathfrak{P})} = \rhd_{IPC,\mathfrak{L}(\mathfrak{P})} =^{(\delta)} \rhd_{HA,\mathfrak{A},\mathfrak{A}} =^{(\delta)} \rhd_{HA,\mathfrak{A},\Sigma} =^{(\delta)} \rhd_{HA^*,\mathfrak{A},\mathfrak{A}}$
$=^{(\delta)} \rhd_{HA^*,\mathfrak{A},\Sigma} =^{(\epsilon)} \rhd_{HA^*,\top,\mathfrak{A}} =^{(\epsilon)} \rhd_{HA^*,\top,\Sigma}$.
   - δ) HA* is introduced in section 6. All the identities follow immediately from De Jongh's Completeness Theorem for Σ-substitutions for HA and HA*. See e.g DV[94].
   - ϵ) The identity of $\rhd_{HA^*,\top,\mathfrak{A}}$ and $\rhd_{HA^*,\top,\Sigma}$ with $\rhd_{IPC,\mathfrak{L}(\mathfrak{P})}$, is corollary 5.7 of DV[94]
3) $\rhd_{IPC,EX} =^{(\eta)} \rhd_{IPC,\top,\mathfrak{L}(\mathfrak{P})} =^{(\zeta)} \rhd_{HA,\top,B(\Sigma)}$.
   - η) See 3.8.
   - ζ) This is theorem 6.2 of DV[94].
4) NNIL $=^{(\vartheta)}$ ROB $=^{(\vartheta)}$ f-ROB (mod IPC), and
$\rhd_{IPC,NNIL} =^{(\iota)} \rhd_{HA,\Sigma,\Sigma} =^{(\iota)} \rhd_{HA,\top,\Sigma}$.
   - ϑ) ROB and f-ROB are defined in section 4 below. The result is proved in Visser[85] and by a similar proof in section 5 of this paper. A different proof, due to Johan van Benthem, is contained in Van Benthem[91]. A version of van Benthem's proof is contained in DRVV[94].
   - ι) See section 9.

In the next section we study robust formulas, as a preliminary to the specification of the NNIL-algorithm in section 5.

4 Robust formulas: Consider a $\mathfrak{P}$-models $\mathbb{K}$ and $\mathbb{M}$. We say that:

- $\mathbb{K} \subseteq \mathbb{M} :\Leftrightarrow \exists f{:}K_{\mathbb{K}} \rightarrow K_{\mathbb{M}}$, f is injective and
$\leq_{\mathbb{K}} \circ f \subseteq f \circ \leq_{\mathbb{M}}$ and $\forall k \in K_{\mathbb{K}}, p \in \mathfrak{P}$ $(k \Vdash_{\mathbb{K}} p \Leftrightarrow f(k) \Vdash_{\mathbb{M}} p)$.

Note that, modulo the identification of the elements of K with their f-images in M, '$\mathbb{K} \subseteq \mathbb{M}$' means that $\mathbb{K}$ is a submodel of $\mathbb{M}$.

- $A \in ROB :\Leftrightarrow A$ is robust $:\Leftrightarrow \forall \mathbb{M} (\mathbb{M} \Vdash A \Rightarrow \forall \mathbb{K} \subseteq \mathbb{M}\ \mathbb{K} \Vdash A)$

We will let $\sigma,\sigma',\tau,...$ range over ROB. It is easy to see that NNIL$\subseteq$ROB. In section 5 we will see that modulo IPC-provable equivalence each robust formula is in NNIL.

In this section we will prove that $\rhd_{IPC,ROB}$ is a σ-relation. This result will be our main tool (in section 5) for proving the NNIL-algorithm to be correct. Let's take as a local convention of this section: $\rhd := \rhd_{IPC,ROB}$.

Clearly we have: $\mathfrak{P}\subseteq$ROB and ROB is closed under conjunction. So it follows that $\rhd$ satisfies A1,A2,A3,B2. To verify B1, by 3.5(ii), the following theorem suffices:

4.1 Theorem: ROB is weakly disjunctive.

Proof: Consider any $\sigma\in$ROB. We write σ in disjunctive normal form D_σ as in 2.5.1. Consider any disjunct $C(\Delta)$ of D_σ: here, as in 2.5.1, Δ is a $\subseteq$-minimal SUB(σ)-saturated set with $\sigma\in\Delta$ and $C(\Delta)$ is the conjunction of the atoms and implications in Δ. We claim that $C(\Delta)$ is robust. Consider any models $\mathbb{K}\subseteq\mathbb{M}\models C(\Delta)$. Trivially: $\mathbb{M}\models\Delta$. By the Push Down Lemma 2.3:

$$\text{Glue}(\mathbb{H}_{SUB(\sigma)}[\Delta],\mathbb{M})\models\Delta.$$

Hence:

$$\text{Glue}(\mathbb{H}_{SUB(\sigma)}[\Delta],\mathbb{M})\models\sigma.$$

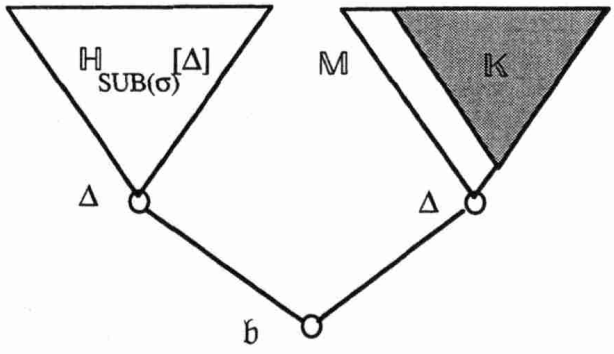
Now clearly:

$$\text{Glue}(\mathbb{H}_{SUB(\sigma)}(\Delta),\mathbb{K}) \subseteq \text{Glue}(\mathbb{H}_{SUB(\sigma)}(\Delta),\mathbb{M}).$$

By the stability of σ we get: $\text{Glue}(\mathbb{H}_{SUB(\sigma)}(\Delta),\mathbb{K})\models\sigma$. Consider

$$\Gamma := \{G\in SUB(\sigma) \mid \text{Glue}(\mathbb{H}_{SUB(\sigma)}(\Delta),\mathbb{K})\models G\}.$$

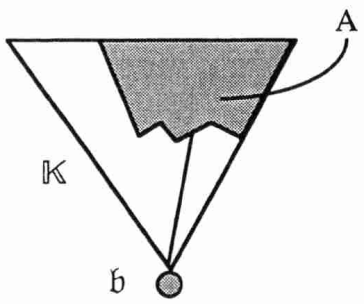
Clearly $\Gamma\subseteq\Delta$ and Γ is SUB(σ)-saturated and $\sigma\in\Gamma$. By the $\subseteq$-minimality of Δ we find: $\Gamma=\Delta$. Hence $\text{Glue}(\mathbb{H}_{SUB(\sigma)}(\Delta),\mathbb{K})\models C(\Delta)$ and so $\mathbb{K}\models C(\Delta)$. Ergo $C(\Delta)$ is robust. ❑



4.2 Theorem: $\rhd$ is closed under B3.

Proof: Let X be a finite set of implications and let $Y:=\{C|(C\to D)\in X\}\cup\{B\}$. Let $A:=\bigwedge X$. We have to show: $(A\to B)\rhd[A]Y$. The proof is by contraposition. Consider any $\sigma\in ROB$ and suppose: $\sigma\not\Vdash[A]Y$. Let $\mathbb{K}=\langle K,\mathfrak{b},\leq,\Vdash,\mathfrak{P}\rangle$ be a rooted model such that $\mathbb{K}\Vdash\sigma$ and $\mathbb{K}\not\Vdash\bigvee[A]Y$, i.e. for all $E\in Y$: $\mathbb{K}\not\Vdash[A]E$.

Let $\mathbb{K}':=\mathbb{K}\{A\}:=\langle K',\mathfrak{b},\leq',\Vdash',\mathfrak{P}\rangle$, be the full submodel of $\mathbb{K}$ on $K':=\{\mathfrak{b}\}\cup\{k'\in K|\ k\Vdash A\}$. (A submodel is *full* if the new ordering relation is the restriction of the old ordering relation to the new worlds.) Note that $\{k'\in K|\ k\Vdash A\}$ is upwards closed and that on $\{k'\in K|\ k\Vdash A\}$ the old and the new forcing coincide. Moreover on this class [A]G is equivalent with G.



4.2.1 **Claim:** For all F: $\mathfrak{b}\Vdash'F\Rightarrow\mathfrak{b}\Vdash[A]F$.

Proof of the claim: The proof is by a simple induction on F. The cases of atoms disjunction and conjunction are trivial. Suppose F is an implication and $\mathfrak{b}\Vdash'F$. Then certainly: for all $k\in K$ $(k\Vdash A\Rightarrow k\Vdash'F)$. Since on the k with $k\Vdash A$, $\Vdash$ and $\Vdash'$ coincide, we find: $\mathfrak{b}\Vdash(A\to F)$, i.e. $\mathfrak{b}\Vdash[A]F$. ❑(Claim)

We return to the main proof. Remember that:

$\mathfrak{b}\Vdash\sigma$ and $\mathfrak{b}\not\Vdash[A]C$ for all C with $(C\to D)\in X$ and $\mathfrak{b}\not\Vdash[A]B$.

We show that $\mathfrak{b}\Vdash'\sigma$ and $\mathfrak{b}\Vdash'A$ and $\mathfrak{b}\not\Vdash'B$.

It is immediate that $\mathfrak{b}\Vdash'\sigma$, since $\mathbb{K}'$ is a submodel of $\mathbb{K}$ and σ is robust.

Remember that A is the conjunction of the $(C\to D)$ in X. So it is sufficient to show that for each $(C\to D)$ in X: $\mathfrak{b}\Vdash'(C\to D)$. Consider any $(C\to D)\in X$ and any $k'\geq'\mathfrak{b}$ with $k'\Vdash'C$. Since $\mathfrak{b}\not\Vdash[A]C$, we have by the claim: $\mathfrak{b}\not\Vdash'C$. So $k'\neq\mathfrak{b}$. But then $k'\Vdash A$, hence $k'\Vdash'A$ and thus $k'\Vdash'(C\to D)$. We may conclude: $k'\Vdash'D$ and hence $\mathfrak{b}\Vdash'(C\to D)$.

From $\mathfrak{b}\not\Vdash[A]B$ and the claim we have immediately that: $\mathfrak{b}\not\Vdash'B$. ❑

All the proofs in this section also work when we replace ROB by f-ROB, the set of formulas preserved by full submodels. Note that ROB$\subseteq$f-ROB. Our result in section 5 will imply: ROB=f-ROB=NNIL (modulo IPC-provable equivalence).

5 The NNIL-algorithm: In this section we produce the algorithm that is the main tool of this paper. The existence of the algorithm establishes the following theorem.

5.1 Theorem: For all A there is an $A^* \in NNIL(PV(A))$ such that for all σ-relations $\rhd$: $A \rhd A^*$ and $A^* \vdash A$.

For convenience we reproduce the defining properties of σ-relations here.

A1 $A \vdash B \Rightarrow A \rhd B$

A2 $A \rhd B$ and $B \rhd C \Rightarrow A \rhd C$

A3 $C \rhd A$ and $C \rhd B \Rightarrow C \rhd (A \wedge B)$

B1 $A \rhd C$ and $B \rhd C \Rightarrow (A \vee B) \rhd C$

B2 $A \rhd B \Rightarrow (p \rightarrow A) \rhd (p \rightarrow B)$

B3 Let X be a finite set of implications and let $Y := \{Cl(C \rightarrow D) \in X\} \cup \{B\}$,. Take $A := \bigwedge X$. Then: $(A \rightarrow B) \rhd [A]Y$

Before proceeding with the proof of 5.1 we interpolate a corollary. Let's write "mod IPC" for "modulo IPC-provable equivalence".

5.2 Corollary

i) Let A and A* be as promised by 5.1. Then we have:
$$A \rhd_{ROB} B \Leftrightarrow A^* \vdash B.$$
It follows that (.)* has the properties promised in 3.2.

ii) NNIL=ROB=f-ROB (mod IPC)

iii) $\rhd_{ROB}$ is the minimal σ-relation. It follows that A1-A3,B1-B3 axiomatizes $\rhd_{ROB}$.

Proof: (i) Suppose $A \rhd_{ROB} B$. Since $A^* \vdash A$, it follows by A1 that $A^* \rhd_{ROB} A$ and, hence, by A2 that $A^* \rhd B$. Since $A^* \in NNIL \subseteq ROB$, we have, by 3.5(v), that $A^* \vdash B$. Conversely if $A^* \vdash B$, then $A^* \rhd_{ROB} B$. Since $\rhd_{ROB}$ is a σ-relation, we have: $A \rhd_{ROB} A^*$ and hence by A2: $A \rhd_{ROB} B$.

(ii) Consider $A \in$ f-ROB. We have $A \rhd_{f\text{-}ROB} A^*$, so by 3.5(v): $A \vdash A^*$. Since also $A^* \vdash A$, we find: $\vdash A \leftrightarrow A^*$. Ergo f-ROB$\subseteq$NNIL (mod IPC). Since obviously NNIL$\subseteq$ROB$\subseteq$f-ROB, we are done.

(iii) Let $\rhd$ be any σ-relation. We have by (i), A1, 5.1, A2:

$$A \rhd_{ROB} B \Rightarrow A^* \vdash B \Rightarrow A^* \rhd B \Rightarrow A \rhd B. \qquad \square$$

5.2(ii) was proved by purely model theoretical means by Johan van Benthem. See his Van Benthem[91] or DRVV[94]. The advantage of van Benthem's proof is its relative simplicity and the fact that the method employed easily generalizes. The advantage of the present method is the extra information it produces, like 5.2(iii) and its usefulness in the arithmetical case, see sections 8, 9 and 10. It is an open question whether van Benthem's proof can be adapted to the arithmetical case.

Proof of 5.1: We introduce an ordinal measure $\mathfrak{o}$ of complexity on formulas as follows:

- $I(D) := \{E \in SUB(D) \mid E \text{ is an implication}\}$,
- $\mathfrak{i}(D) := \max\{|I(E)| + 1 \mid E \in I(D)\}$,
- $\mathfrak{c}(D) :=$ the number of occurrences of logical connectives in D,
- $\mathfrak{o}(D) := \omega.\mathfrak{i}(D) + \mathfrak{c}(D)$.

Note that we count *occurrences* of connectives for $\mathfrak{c}$ and *types* of implications, not occurrences, for $\mathfrak{i}$!

We say that an occurence of E in D is an *outer occurrence* if this occurrence is not in the scope of an implication.

We prove 5.1 by induction on $\mathfrak{o}$. Consider any σ-relation $\rhd$. We will exhibit the properties of $\rhd$ that are used between square brackets.

α **Atoms:** [A1] Suppose A is an atom. Take $A^* := A$.

β **Conjunction:** [A1,A2,A3]: Suppose $A = (B \wedge C)$. Clearly $\mathfrak{o}(B) < \mathfrak{o}(A)$ and $\mathfrak{o}(C) < \mathfrak{o}(A)$. Take $A^* := B^* \wedge C^*$. Clearly $A^* \in NNIL(PV(A))$. The verification of the desired properties is trivial.

γ **Disjunction:** [A1,A2,B1] Suppose $A = (B \vee C)$. Clearly $\mathfrak{o}(B) < \mathfrak{o}(A)$ and $\mathfrak{o}(C) < \mathfrak{o}(A)$. Take $A^* := B^* \vee C^*$. Clearly $A^* \in NNIL$. The verification of the desired properties is trivial.

δ **Implication:** [A1,A2,A3,B2,B3] Suppose $A = (B \rightarrow C)$. We split into several cases.

δ1 Outer conjunction in the consequent: [A1,A2,A3] Suppose C has an outer ocurrence of a formula $D \wedge E$. Pick any J(q) be such that:

- q is a fresh variable,
- q occurs precisely once in J,
- q is not in the scope of an implication in J,
- $C=J[q:=D \wedge E]$.

Let $C_1:=J[q:=D]$, $C_2:=J[q:=E]$. As is easily seen C is IPC-provably equivalent to $C_1 \wedge C_2$. Let $A_1:=(B \rightarrow C_1)$ and $A_2:= (B \rightarrow C_2)$. Clearly A is IPC-provably equivalent to $A_1 \wedge A_2$. We prove: $\mathfrak{o}(A_i)<A$ for i=0,1. Since it is clear that $\mathfrak{c}(A_i)<\mathfrak{c}(A)$, it is sufficient to show that $\mathfrak{i}(A_i) \leq \mathfrak{i}(A)$. Since A and the A_i are implications we have to show that: $|I(A_i)| \leq |I(A)|$. We treat the case that i=1. It is sufficient to construct an injective mapping from $I(A_1)$ to I(A). Consider any implication F in $I(A_1)$. If $F=A_1$, we map F to A. Otherwise $F \in I(B)$ or $F \in I(J)$ or $F \in I(D)$ (since q does not occur in the scope of an implication). In all three cases we can map F to itself. Since A_1 cannot be in I(B) or I(J) or I(D), our mapping is injective. The case that i=2 is similar. Set $A^* := (A_1^* \wedge A_2^*)$.

δ2 Outer disjunction in the antecedent: [A1,A2,A3] This case is completely analogous to the previous one.

If A has no outer disjunction in the antecedent and no outer conjunction in the consequent, then B is a conjunction of atoms and implications and C is a disjunction of atoms and implications. It is easy to see, that applications of associativity, commutativity and idempotency to the conjunction in the antecedent or to the disjunction in the consequent do not raise $\mathfrak{o}$. So we can safely write: $B=\bigwedge X$ and $C=\bigvee Y$, where X and Y are finite sets of atoms or implications. This leads us to the following case.

δ3 B is a conjunction of atoms and implications, C is a disjunction of atoms and implications [A1,A2,A3,B2,B3]

δ3.1 X contains an atom: [A1,A2,B2]

δ3.1.1 X contains a propositional variable, say p: [A1,A2,B2] Consider: $D:=\bigwedge(X/\{p\})$ and $E:=(D \rightarrow C)$. Clearly $\vdash A \leftrightarrow (p \rightarrow E)$ and $\mathfrak{o}(E)<\mathfrak{o}(A)$. Put $A^*:=(p \rightarrow E^*)$. Evidently $(p \rightarrow E^*)$ is in NNIL(PV(A)). We have $E^* \vdash E$ by the Induction Hypothesis and hence: $(p \rightarrow E^*) \vdash (p \rightarrow E) \vdash A$. We have $E \rhd E^*$ by the Induction Hypothesis. It follows by B2 that: $(p \rightarrow E) \rhd (p \rightarrow E^*)$ From $A \vdash (p \rightarrow E)$, we have by A1: $A \rhd (p \rightarrow E)$. Hence by A2: $A \rhd (p \rightarrow E^*)$.

δ3.1.2 **X contains $\top$:** [A1,A2] Left to the reader.

δ3.1.3 **X contains $\bot$:** [A1] Left to the reader.

δ3.2 **X contains no atoms:** [A1,A2,A3,B3] This case -the last one- is the truly difficult one. To motivate it, let's solve the difficulties one by one. Let's first look at an example.

Example 1: Consider $(p\to q)\to r$. B3 gives us: $((p\to q)\to r)\rhd(p\vee r)$. However, we do not have: $(p\vee r)\vdash((p\to q)\to r)$. We can repair this by noting that $(p\to q)\to r)\vdash(q\to r)$ and $(p\to q)\wedge(p\vee r)\vdash((p\to q)\to r)$. So the full solution of our example is as follows.

We have: $((p\to q)\to r)\rhd((q\to r)\wedge(p\vee r))$, by:

a)	$((p\to q)\to r)\rhd(p\vee r)$	B3
b)	$((p\to q)\to r)\vdash(q\to r)$	IPC
c)	$((p\to q)\to r)\rhd(q\to r)$	A1
d)	$((p\to q)\to r)\rhd((q\to r)\wedge(p\vee r))$	a,c,A3

Moreover: $((q\to r)\wedge(p\vee r))\vdash((p\to q)\to r)$ and $((q\to r)\wedge(p\vee r))\in NNIL(PV((p\to q)\to r))$.

○

We implement this idea for the general case. There will be a problem with $\mathfrak{o}$, but we will postpone its discussion until we run into it. A is of the form $B\to C$, where B is a conjunction of a finite set of implications X and C is a disjunction of a finite set of atoms or implications Y. For any $D:=(E\to F)\in X$, let:

- $B{\downarrow}D := \bigwedge((X/\{D\})\cup\{F\})$.

Clearly and $\mathfrak{o}((B{\downarrow}D)\to C) < \mathfrak{o}(A)$.

Let $Z:=\{E|(E\to F)\in X\}\cup\{C\}$. Put: $A_0 := \bigvee[B]Z$. We will show that our problem reduces to the question whether $A_0{}^*$ exists. So for the moment we pretend it does. We have:

a)	$A\rhd A_0$	B3
a1)	$A_0\rhd A_0{}^*$	ASS
a2)	$A\rhd A_0{}^*$	a,a1,A2
b)	for all $D\in X$: $A\vdash((B{\downarrow}D)\to C)$	IPC
c)	for all $D\in X$: $A\rhd((B{\downarrow}D)\to C)$	A1
c1)	for all $D\in X$: $((B{\downarrow}D)\to C)\rhd((B{\downarrow}D)\to C)^*$	IH
c2)	for all $D\in X$: $A\rhd((B{\downarrow}D)\to C)^*$	c,c1,A2
d)	$A\rhd\bigwedge\{((B{\downarrow}D)\to C)^*\|D\in X\}\wedge A_0{}^*$	a2,c2,A3.

It is clear that $\bigwedge\{((B\downarrow D)\rightarrow C)^* | D\in X\}\wedge A_0^* \in NNIL(PV(A))$. We show that:

$\bigwedge\{((B\downarrow D)\rightarrow C)^* | D\in X\}\wedge A_0^* \vdash A$.

It is sufficient to show:

$\bigwedge\{((B\downarrow D)\rightarrow C) | D\in X\}\wedge [B]E \vdash A$ for each $E\in Z$.

In case E=C, we are immediately done by: $[B]C \vdash B\rightarrow C$. Suppose $(E\rightarrow F)\in X$ for some F. Reason in IPC.

Suppose $\bigwedge\{((B\downarrow D)\rightarrow C) | D\in X\}$, [B]E and B. We want to derive C. We have: $((\bigwedge(X/\{E\rightarrow F\})\wedge F)\rightarrow C)$, [B]E and B. From B we find $\bigwedge(X/\{E\rightarrow F\})$ and $(E\rightarrow F)$. From B and [B]E, we derive E. From E and $(E\rightarrow F)$ we get F. Finally we infer from $((\bigwedge(X/\{E\rightarrow F\})\wedge F)\rightarrow C)$, $\bigwedge(X/\{E\rightarrow F\})$ and F the desired conclusion C.

So the only thing left to do is to show that A_0^* exists. If $i(A_0)=0$, we are easily done. Suppose $i(A_0)>0$. If for each E in Z with $i(E)>0$, we would have: $i([B]E)<i(A)$, it would follow that $i(A_0)<i(A)$. Hence we would be done by the Induction Hypothesis.

We study some examples. These examples show that we cannot generally hope to get $i([B]E)<i(A)$. The examples will, however, suggest a way around the problem: we produce a logical equivalent, say Q, of [B]E, such that $i(Q)<i(A)$. So we replace A_0 by the disjunction R of the Q's, which is logically equivalent and has $i(R)<i(A_0)$. Put $A_0^*:=R^*$.

Example 2: Consider $(p\rightarrow q)\rightarrow(r\rightarrow s)$. Take $E:=C=(r\rightarrow s)$. We have:

$[B]E = [p\rightarrow q](r\rightarrow s) = (p\rightarrow q)\rightarrow(r\rightarrow s)$.

So no simplification is obtained. However [B]E is IPC-provably equivalent to: $((r\wedge(p\rightarrow q))\rightarrow s)$, which has lower i. By A1,A2 we can put $([B]E)^*:=((r\wedge(p\rightarrow q))\rightarrow s)^*$.

We could have applied the reduction before going to [B]E. We did not choose to do so for reasons of uniformity of treatement. ○

Example 3: Consider $(((p\rightarrow q)\rightarrow r)\rightarrow s)$. Take $E:=(p\rightarrow q)$. We find:

$[B]E = [(p\rightarrow q)\rightarrow r](p\rightarrow q) = (((p\rightarrow q)\rightarrow r)\rightarrow(p\rightarrow q))$.

[B]E is IPC-provably equivalent to $((p\wedge(q\rightarrow r))\rightarrow q)$, which has lower i. ○

Consider [B]E for $E\in Z=\{E | (E\rightarrow F)\in X\}\cup\{C\}$. [B]E is the result over replacing outer implications $(G\rightarrow H)$ of E by $(B\rightarrow(G\rightarrow H))$. If there is no outer implication in E, we find: $[B]E = E$ and $i([B]E) = 0$. In this case [B]E is implication free and hence a fortiori in NNIL. We can put $([B]E)^*:=E$. Suppose E has an outer implication.

Consider any outer implication $(G\to H)$ of E. We first show: $\dot{\imath}(B\to(G\to H))\leq\dot{\imath}(B\to C)$. We define an injection from $I(B\to(G\to H))$ to $I(B\to C)$. Any implication occurring in $I(B)$ or $I(G\to H)$ can be mapped to itself in $I(B\to C)$. None of these implications has as image $(B\to C)$, since $(G\to H)\in I(B)\cup I(C)$. So we can send $(B\to(G\to H))$ to $(B\to C)$.

The next step is to replace $(B\to(G\to H))$ by an IPC-provably equivalent formula K with lower $\dot{\imath}$. Let B' be the result of replacing all occurrences of $(G\to H)$ in B by H and let: $K := ((G\wedge B')\to H)$. Clearly K is provably equivalent to $(B\to(G\to H))$. We show that $\dot{\imath}(K)<\dot{\imath}(B\to(G\to H))$.

We define a non-surjective injection from I(K) to $I(B\to(G\to H))$. Consider an implication M in I(K). If M=K we send it to $(B\to(G\to H))$. Suppose $M\neq K$, i.e. $M\in I(B')\cup I(G)\cup I(H)$. A subformula N of $I(B)\cup I(G)\cup I(H)$ is a *predecessor* of M if M is the result of replacing all occurrences of $(G\to H)$ in N by H. Clearly M has at least one predecessor and any predecessor of M must be an implication. Send M to one of its shortest predecessors. Clearly our function is injective, since two different implications cannot share a predecessor. Finally $(G\to H)$ cannot be in the range of our injection. If it were, we would have M=H, but then H would be a shorter predecessor. A contradiction.

Replace every outer implication in [B]E by an equivalent with lower $\dot{\imath}$. The result is the desired Q. ❑

5.3 Remarks: The algorithm given with our proof is non-deterministically specified. However by 5.2 the result is unique modulo provable equivalence. I didn't try to make the algorithm optimally quick. In practice the best thing to do is: follow the main line of the algorithm, but, locally, apply ad hoc simplifications.

In Visser[85] a simple adaptation of the NNIL algorithm is given for $\rhd_{IPC,\top}$. Here the algorithm computes not a value in NNIL, but a value in $\{\top,\bot\}$. Thus we obtain an algorith that checks for provability. It is easy to use the present algorithm for this purpose too, since there is a p-time algorithm to decide IPC-provability of $NNIL_0$ formulas, i.e. formulas with only propositional variables in front of arrows. The outputs of our algorithm are in $NNIL_0$. It has been shown by Richard Statman (see Statman[79]) that checking whether a formula is IPC-provable is p-space complete. This puts an absolute bound on what an algorithm like ours can do.

5.4 **Sample computations** (using some obvious short-cuts)

$(\neg\neg p \to p) \to (p \vee \neg p) \equiv$

$(p \to (p \vee \neg p)) \wedge (([\neg\neg p \to p]\neg\neg p) \vee ([\neg\neg p \to p]p) \vee ([\neg\neg p \to p]\neg p)) \equiv$

$\neg\neg p \vee p \vee \neg p \equiv$

$((\bot \to p) \wedge ([\neg p]p \vee [\neg p]\bot)) \vee p \vee \neg p \equiv$

$p \vee \neg p.$

$((p \to q) \to r) \to s \equiv$

$(r \to s) \wedge (([(p \to q) \to r](p \to q)) \vee [(p \to q) \to r]s) \equiv$

$(r \to s) \wedge (((p \wedge (q \to r)) \to q) \vee s) \equiv$

$(r \to s) \wedge ((p \to ((q \to r) \to q)) \vee s) \equiv$

$(r \to s) \wedge ((p \to ((r \to q) \wedge ([q \to r]q \vee [q \to r]q))) \vee s) \equiv$

$(r \to s) \wedge ((p \to q) \vee s).$

6 Basic facts and notations in Arithmetic

6.1 **Arithmetical Theories:** The arithmetical theories T considered in this paper are RE theories in $\mathfrak{A}$, the language of HA. These theories all extend i-IΔ_0+Exp: the constructive theory of Δ_0-induction with the axiom expressing that the exponentiation function is total. i-IΔ_0+Exp is finitely axiomatizable; we stipulate that a fixed finite axiomatization is employed. We will use $\Box_T$ for the formalization of provability in T. Suppose A is a formula. $\Box_T$A means $\mathrm{Prov}_T(t(\mathbf{x}))$, where Prov_T is the arithmetization of the provability predicate of T and where $t(\mathbf{x})$ is the term 'the Gödelnumber of the result of substituting the Gödelnumbers of the numerals of the **x**'s for the variables in A'.

We illustrate the above by an example. We suppose Gödelnumbers are assigned as follows:

(	11
)	12
=	15
S	8
0	3
+	19

$*$ is the arithmetization of the syntactical operation of concatenation. We use underlining for our external numeral function. *num* is the arithmetization of the numeral function. We have e.g.: HA $\vdash$ num(3) = $\underline{8}*\underline{8}*\underline{8}*\underline{3}$. And:

$\Box_T(x=x)$ means: $\mathrm{Prov}_T(\underline{11}*\mathrm{num}(x)*\underline{15}*\mathrm{num}(x)*\underline{12})$.

Our notational convention evidently introduces a scope ambiguity. What is $\Box_T((x+y)=z)$ going to mean?

a) $\mathrm{Prov}_T(\underline{11}*\mathrm{num}(x+y)*\underline{15}*\mathrm{num}(z)*\underline{12})$ (wide scope)

or:

b) $\mathrm{Prov}_T(\underline{11}*\underline{11}*\mathrm{num}(x)*\underline{19}*\mathrm{num}(y)*\underline{12}*\underline{15}*\mathrm{num}(z)*\underline{12})$ (small scope)

Fortunately by standard metamathematical results, we know that as long as the terms we employ stand for T-provably total recursive functions the different readings are provably equivalent. So (a) and (b) are provably equivalent. In this paper we will only employ terms for primitive recursive functions, so the ambiguity is harmless.

We write T_n for the theory axiomatized by the finitely many axioms of i-IΔ_0+Exp, plus the axioms of T, which are smaller than n in the standard Gödelnumbering. We write $\mathrm{Prov}_{T,n}$ for the formalization of provability in T_n. We consider $\mathrm{Prov}_{T,n}$ as a form of *restricted provability* in T. The following well-known fact is quite important:

6.1.1 Fact: Suppose T is an RE extension of HA in the language of HA. Then T is essentially reflexive (verifiably in HA). I.e. we have:

For all n and for all $\mathfrak{A}$-formulas A with free variables $\mathbf{x}$: $T\vdash \forall \mathbf{x}\,(\Box_{T,n}A \rightarrow A)$.

And (using UC for: 'the universal closure of') even:

$HA\vdash \forall x \forall A \in \mathrm{FOR}_{\mathfrak{A}}\, \Box_T \mathrm{UC}(\Box_{T,x}A \rightarrow A)$.

Proofsketch: The proof is roughly as follows. Ordinary cut-elimination for constructive predicate logic (or normalization in case we have a natural deduction system) can be formalized in HA. Reason in HA. Let a number x and a formula A be given. Introduce a measure of complexity on arithmetical formulas that counts both the depth of quantifiers and of implications. Find y such that both the axioms of T_x and A have complexity <y. We can construct a truthpredicate True_y for formulas of complexity <y. We have: $\Box_T \mathrm{UC}(\mathrm{True}_y A \rightarrow A)$. Reason inside $\Box_T$. Suppose we have $\Box_{T,x}A$. By cut-elimination we can find a T_x-proof p of A in which only formulas of complexity <y occur. We now prove by induction on the subproofs of p, that all subconclusions of p are True_y. So A is True_y. Hence we find A. ❑

6.2 A brief introduction to HA*: In this section we describe the theory HA*. This theory was introduced in Visser[82]. HA* is to Beeson's fp-realizability (Beeson[75]) as Troelstra's $\mathrm{HA+ECT}_0$ is to Kleene's r-realizability. This means that for a suitable variant of fp-realizability HA* is the set of sentences such that their fp-translations are provable in HA. The natural way to define HA* is by a fixed point construction as: HA plus the *Completeness Principle for HA**. (Here it is essential that the construction is verifiable in HA, see below.) The Completeness Principle can be

viewed as an arithmetically interpreted modal principle. The Completeness Principle viewed modally is:

$\mathfrak{C}$ $\quad \vdash A \rightarrow \Box A$

The Completeness Principle for a specific theory T is:

$\mathfrak{C}[T]$ $\quad \vdash A \rightarrow \Box_T A$.

We have: $HA^* = HA + \mathfrak{C}[HA^*]$.

We briefly review some of the results of Visser[82] and DV[94].

- Let **A** be the smallest class closed under atoms and all connectives except implication, satisfying: $A \in \Sigma_1$ and $B \in \mathbf{A} \Rightarrow (A \rightarrow B) \in \mathbf{A}$. Note that modulo provable equivalence in HA all formulas of the classical arithmetical hierarchy in their standard form are in **A**. HA^* is conservative w.r.t. **A** over HA. Note that $NNIL_0(\Sigma) \subseteq \mathbf{A}$.
- There are infinitely many incomparable T with $T=HA+\mathfrak{C}[T]$. However if $T=HA+\mathfrak{C}[T]$ *verifiably in HA*, then $T=HA^*$.
- Let KLS:=Kreisel-Lacombe-Shoenfield's Theorem on the continuity of the effective operations. We have: $HA^* \vdash KLS \rightarrow \Box_{HA^*}\bot$. This immediately gives Beeson's result that $HA \nvdash KLS$ (see Beeson[75]).
- Every prime RE Heyting algebra $\mathfrak{H}$ can be embedded into the Heyting algebra of HA*. This mapping is primitive recursive in the enumeration of the generators of $\mathfrak{H}$ w.r.t which $\mathfrak{H}$ is RE. The mapping sends the generators to Σ-sentences. (See DV[94]).

We insert some basic facts on the provability logic of HA*. These materials will be only needed in section 10. Clearly, HA* satisfies the Löb conditions.

L1 $\quad \vdash A \Rightarrow \vdash \Box A$

L2 $\quad \vdash \Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$

L3 $\quad \vdash \Box A \rightarrow \Box\Box A$

L4 $\quad \vdash \Box(\Box A \rightarrow A) \rightarrow \Box A$

i-K is given by IPC+L1,L2. i-L is i-K+L3,L4. We write i-K{P} for the extension of i-K with some principle P. We write '$\mathfrak{C}[HA^*]$' in the modal contexts as C. Note that i-L{C} is valid for provability interpretations in HA^*.

A principle closely connected to C is the Strong Löb Principle:

SL $\quad \vdash (\Box A \rightarrow A) \rightarrow A$

As a special case of SL we have: $\vdash \neg\neg\Box\bot$.

6.2.1 Fact: i-L{C} is interderivable with i-K{SL}.

Proof: L4 is immediate from SL. "i-K{SL}$\vdash$C":

$\vdash A \rightarrow (\Box(A\wedge\Box A) \rightarrow (A\wedge\Box A))$

$\rightarrow A\wedge\Box A$

$\rightarrow \Box A.$

"i-L{C}$\vdash$SL":

$\vdash (\Box A\rightarrow A) \rightarrow (\Box A\rightarrow A)\wedge \Box(\Box A\rightarrow A)$

$\rightarrow (\Box A\rightarrow A)\wedge \Box A$

$\rightarrow A.$ ❑

As a preliminary for section 10 we study the closed fragment of i-L{C}. A formula of the modal language is closed if it contains no propositional variables. We define: $\Box^0\bot := \bot$, $\Box^{n+1}\bot := \Box\Box^n\bot$, $\Box^\omega\bot := \top$.

6.2.2 The Closed Fragment of i-L{C}: For every closed formula A there is an $\alpha =: \alpha^*(A)$, such that $\vdash A \leftrightarrow \Box^\alpha\bot$. It is easily seen that $\alpha^*(A)$ is unique.

Proof: The proof is by induction on A. We have:

$\vdash \top \leftrightarrow \Box^\omega\bot, \vdash \bot \leftrightarrow \Box^0\bot$

$\vdash (\Box^\alpha\bot\wedge\Box^\beta\bot) \leftrightarrow \Box^{\min(\alpha,\beta)}\bot$

$\vdash (\Box^\alpha\bot\vee\Box^\beta\bot) \leftrightarrow \Box^{\max(\alpha,\beta)}\bot$

$\vdash (\Box^\alpha\bot\rightarrow\Box^\beta\bot) \leftrightarrow \Box^{\alpha\rightarrow\beta}\bot$, where $(\alpha\rightarrow\beta) := \top$ if $\alpha\leq\beta$ and $(\alpha\rightarrow\beta) := \beta$ if $\beta<\alpha$.

Note that $\min(\alpha,\gamma)\leq\beta \Leftrightarrow \gamma\leq(\alpha\rightarrow\beta)$.

$\vdash \Box\Box^\alpha\bot \leftrightarrow \Box^{1+\alpha}\bot$ ❑

7 Digression: assorted facts about Σ-preservativity: In this section we pause to look in some more detail at the notion of Σ-preservativity. We will transfer some classical results about Π-conservativity to Σ-preservativity and we formulate some principles of 'preservativity logic'. *The only result of this section that is used in the rest of the paper is 7.1.*

Consider any consistent RE theory T, extending HA, in the language of HA. We introduce some notions, closely related to Σ-preservativity for T. Define (suppressing the index for T):

Provable Deductive Consequence	$A \rhd_{pdc} B$	$:\leftrightarrow \Box(\Box A \rightarrow \Box B)$
Uniform Deductive Consequence	$A \rhd_{udc} B$	$:\leftrightarrow \forall x \exists y\, (\Box_x A \rightarrow \Box_y B)$
Strong Uniform Deductive Consequence	$A \rhd_{sudc} B$	$:\leftrightarrow \forall x\, (\Box_x A \rightarrow \Box_x B)$
Uniform Provably Deductive Consequence	$A \rhd_{updc} B$	$:\leftrightarrow \forall x \exists y\, \Box(\Box_x A \rightarrow \Box_y B)$

We start with some equivalences and derivabilities.

7.1 Orey-Hájek for Σ-preservativity and Uniform Provable Deductive Consequence: T proves that the following are equivalent:
(i) $A \rhd_\Sigma B$, (ii) $\forall x\, \Box(\Box_x A \rightarrow B)$, (iii) $A \rhd_{updc} B$.

Proof: Reason in T. "(i)$\rightarrow$(ii)" Suppose $A \rhd_\Sigma B$ and consider any x. We have $\Box(\Box_x A \rightarrow A)$ and $(\Box_x A) \in \Sigma$, hence $\Box(\Box_x A \rightarrow B)$.

"(ii)$\rightarrow$(iii)" From $\Box(\Box_x A \rightarrow B)$, we have that for some y $\Box_y(\Box_x A \rightarrow B)$ and hence $\Box\Box_y(\Box_x A \rightarrow B)$. Ergo: $\Box(\Box_y \Box_x A \rightarrow \Box_y B)$. Clearly for any u: $\Box(\Box_u A \rightarrow \Box_y \Box_u A)$. We may conclude: $\Box(\Box_x A \rightarrow \Box_y B)$. Hence: $A \rhd_{updc} B$.

"(i)$\leftarrow$(iii)" Suppose $A \rhd_{updc} B$ and $\Box(S \rightarrow A)$. It follows that for some x: $\Box\Box_x(S \rightarrow A)$ Moreover: $\Box(S \rightarrow \Box_x S)$. So $\Box(S \rightarrow \Box_x A)$. Hence for some y: $\Box(S \rightarrow \Box_y B)$. Ergo by reflection: $\Box(S \rightarrow B)$. ❑

7.2 Fact: We have: $T \vdash A \rhd_\Sigma B \rightarrow A \rhd_{udc} B$.

Proof: Reason in T. Suppose $A \rhd_\Sigma B$. Consider any x. For some y we have: $\Box_y(\Box_x A \rightarrow B)$. Suppose $\Box_x A$. Clearly for any u: $\Box_u A \rightarrow \Box_y \Box_u A$. It follows that $\Box_y B$.
❑

7.3 Orey Hájek in HA*: We have: $HA^* \vdash A \rhd_{HA^*,\Sigma} B \leftrightarrow A \rhd_{udc,HA^*} B$.

Proof: Immediate by the principle $\mathfrak{C}$, 7.1 and 7.2. ❑

7.4 Fact: $HA^* \vdash (\Box_{HA^*} A \rightarrow B) \rightarrow A \rhd_{HA^*,\Sigma} B$.

Proof:

$$\begin{aligned} \vdash (\Box A \to B) &\to \forall x(\Box_x A \to B) \\ &\to \forall x\, \Box(\Box_x A \to B) \\ &\to A \rhd_{\Sigma} B. \end{aligned}$$

❑

7.5 Modal logic for Σ-preservativity: Consider the language of modal propositional logic with a unary modal operator $\Box$ and a binary operator $\rhd$. We define arithmetical interpretations in the language of HA in the usual manner, interpreting $\Box$ as provability in HA and $\rhd$ as Σ-preservativity over HA. We state the principles valid in HA for this logic, known at present. With the exception of $\Sigma 4$ and $\Sigma 8$ these principles are the duals of the principles of the interpretability logic ILM.

L1 $\quad \vdash A \Rightarrow \vdash \Box A$

L2 $\quad \vdash \Box(A \to B) \to (\Box A \to \Box B)$

L3 $\quad \vdash \Box A \to \Box\Box A$

L4 $\quad \vdash \Box(\Box A \to A) \to \Box A$

$\Sigma 1$ $\quad \vdash \Box(A \to B) \to A \rhd B$

$\Sigma 2$ $\quad \vdash A \rhd B \wedge B \rhd C \to A \rhd C$

$\Sigma 3$ $\quad \vdash C \rhd A \wedge C \rhd B \to C \rhd (A \wedge B)$

$\Sigma 4$ $\quad \vdash A \rhd C \wedge B \rhd C \to (A \vee B) \rhd C$

$\Sigma 5$ $\quad \vdash A \rhd B \to (\Box A \to \Box B)$

$\Sigma 6$ $\quad \vdash A \rhd \Box A$

$\Sigma 7$ $\quad \vdash A \rhd B \to (\Box C \to A) \rhd (\Box C \to B)$

$\Sigma 8$ $\quad$ Let X be a finite set of implications and let $Y := \{Cl(C \to D) \in X\} \cup \{B\}$. Take $A := \bigwedge X$. Then: $\vdash (A \to B) \rhd \{A\}Y$

Here $(A \to B) \rhd \{A\}Y$ is short for: $(A \to B) \rhd \bigvee \{A\}Y$ and $\{C\}D$ is defined as follows:

- $\{C\}p := (C \to p)$, $\{C\}\bot := \bot$, $\{C\}\top := \top$,
- $\{C\}\Box E := \Box E$, $\{C\}(E \rhd F) = (C \to (E \rhd F))$, $\{C\}(E \to F) := (C \to (E \to F))$,
- $\{C\}(.)$ commutes with $\wedge$ and $\vee$.

The verification of L1-L4, $\Sigma 1$-$\Sigma 3$, $\Sigma 4$-$\Sigma 7$ is routine. For $\Sigma 4$ see 8.1 and for $\Sigma 8$ see 8.2.

From our principles e.g. Leivant's principle can be derived. (This is one of the Stellingen of Leivant[75].)

Le $\quad \vdash \Box(A \vee B) \to \Box(A \vee \Box B)$

We leave this as an exercise to the reader.

It is open whether our axioms are complete.

7.6 Constructive versions of some results of Lindström and Svejdar

In his classical paper Švejdar[83], Vítězslav Švejdar studies logics for interpretability, conservativity and Rosser-orderings. Note that while *the study of methods* is Švejdar's, *the methods studied* are for a large part both Lindström's and Švejdar's.

The definitions of witness comparisons between sentences are as follows:

- $\exists xAx \leq_{ex} \exists yBy :\Leftrightarrow \exists x\,(Ax \wedge \forall y{<}x\ \neg By)$
- $\exists xAx <_{ex} \exists yBy :\Leftrightarrow \exists x\,(Ax \wedge \forall y{\leq}x\ \neg By)$.

If B is decidable these definitions are constructively as useful as classically. However if we allow non-decidable B, the use of negation is rather heavy. I predict that more experimentation will reveal that the following notions are more useful and more pleasant to work with:

- $\exists xAx \leq_{un} \exists yBy :\Leftrightarrow \forall y\,(By \rightarrow \exists x{\leq}y\ Ax)$
- $\exists xAx <_{un} \exists yBy :\Leftrightarrow \forall y\,(By \rightarrow \exists x{<}y\ Ax)$.

Classically we have:

$\exists xAx \leq_{un} \exists yBy$ is provably equivalent to the negation of $\exists xBx <_{ex} \exists yAy$,
$\exists xAx <_{un} \exists yBy$ is provably equivalent to the negation of $\exists xBx \leq_{ex} \exists yAy$,
$\exists xAx <_{ex} \exists yBy$ is provably equivalent to the negation of $\exists xBx \leq_{un} \exists yAy$,
$\exists xAx <_{ex} \exists yBy$ is provably equivalent to the negation of $\exists xBx \leq_{un} \exists yAy$.

Constructively all these connections fail.

We *do* have:

O1 $\vdash \exists xAx \leq_{ex} \exists yBy \rightarrow \exists xAx \leq_{un} \exists yBy$,

O2 $\vdash \exists xAx <_{ex} \exists yBy \rightarrow \exists xAx <_{un} \exists yBy$.

$\leq_{ex}$ and $<_{ex}$ are provably transitive. We have by a 'downwards induction':

O3 $\vdash \neg(\exists xAx <_{ex} \exists xAx)$.

But we do not even have: $\vdash \exists xAx \rightarrow \exists xAx \leq_{ex} \exists xAx$, since this expresses precisely the minimum principle for A, which is constructively not generally valid. (It is valid if A is decidable.)

We have:

O4 $\leq_{un}$ provably satisfies the axioms of weak partial ordering,

O5 $<_{un}$ is provably transitive,

O6 $\vdash (\exists xAx <_{un} \exists xAx) \rightarrow \forall x \neg Ax$,

O7 $\vdash \exists xAx <_{un} \exists yBy \rightarrow \exists xAx \leq_{un} \exists yBy$,

O8 $\vdash \exists xAx \leq_{un} \exists yBy <_{un} \exists zCz \rightarrow \exists xAx <_{un} \exists zCz$,

O9 $\vdash \exists xAx <_{un} \exists yBy \leq_{un} \exists zCz \rightarrow \exists xAx <_{un} \exists zCz$.

Let $\Box^*A :\leftrightarrow \exists x\, \Box_x A$. $\Box_x$ is persistent, i.e. $\vdash (x<y \wedge \Box_x A) \rightarrow \Box_y A$. So we have:

$\vdash \Box^*A <_{ex} \Box^*B \leftrightarrow \exists x\, (\Box_x A \wedge \neg\Box_x B)$,

$\vdash \Box^*A \leq_{un} \Box^*B \leftrightarrow \forall x\, (\Box_x B \rightarrow \Box_x A) \leftrightarrow B \rhd_{sudc} A$.

7.6.1 Svejdar Principles: We give some constructive principles in the style of Švejdar [83]. The superiority of (iii) over (ii) is one of the arguments in favour of the universal witness comparison relation.

i) $\vdash B \rhd \sum C \rightarrow A \rhd \sum((\Box^*B \leq_{un} \Box^*A) \rightarrow C)$

ii) $\vdash A \rhd \sum(\neg B \rightarrow \Box^*A <_{ex} \Box^*B)$

iii) $\vdash A \rhd \sum((B \rightarrow \Box^*A <_{un} \Box^*B) \rightarrow \Box^*A <_{un} \Box^*B)$

iv) $\vdash A \rhd \sum((B \rightarrow \Box^*A \leq_{un} \Box^*B) \rightarrow \Box^*A \leq_{un} \Box^*B)$

Proof: Reason in T

i) Suppose $B \rhd \sum C$. For any x:

$$\Box(\Box_x A \wedge \Box^*B \leq_{un} \Box^*A \rightarrow \Box_x B \rightarrow C).$$

ii) $$\Box(\Box_x A \wedge \neg B \rightarrow \Box_x A \wedge \neg\Box_x B \rightarrow \Box^*A <_{ex} \Box^*B).$$

iii) Reason in T. Consider x. Reason inside the $\Box$. Consider y. Suppose: $\Box_x A$, $(B \rightarrow \Box^*A <_{un} \Box^*B)$ and $\Box_y B$. We have to show: $\exists z<y\, \Box_z A$. If $y \leq x$, then $\Box_x B$ and hence B. So: $\Box^*A <_{un} \Box^*B$. We may conclude $\exists z<y\, \Box_z A$. If $y>x$, we get $\exists z<y\, \Box_z A$ from our assumption $\Box_x A$.

iv) By a minor modification of the reasoning under (iii). ❑

7.6.2 Consequence: $\vdash A \rhd \sum(\Box^*B \leq_{un} \Box^*A \rightarrow B)$.

A version of the Feferman predicate can be defined as follows:

- $\Delta A :\Leftrightarrow \Box^*A <_{ex} \Box^*\bot$.

Note that e.g. $\vdash \neg\Delta\bot$.

7.6.3 Fact: $\vdash A \rhd \sum \Delta A$.

Proof: Substitute $\bot$ for B in 7.6.1(ii). ❑

7.6.4 Orey Sentence: Let $\rhd^*$ be $\prod$-conservativity or interpretability over T (in the classical context). A classical Orey sentence is a sentence G such that:

$\top \rhd^* G$ and $\top \rhd^* \neg G$.

The existence of an Orey-sentence shows the non-uniqueness of extension via interpretation or Π-conservative extension. Its existence also shows that the right hand side of Π-conservativity or interpretability for consistent T cannot be closed under conjunction. By duality an Orey sentence for Σ-preservativity (in the constructive context), should satisfy:

$G \rhd_\Sigma \bot$ and $\neg G \rhd_\Sigma \bot$.

Disanalogously it does *not* generally follow that the left hand side of Σ-preservativity is not closed under disjunction: in fact in the case of HA it is closed under disjunction. See 8.1.

Let G be such that $T \vdash G \leftrightarrow \neg \Delta G$. On the one hand: $G \rhd_\Sigma \Delta G$ and hence $G \rhd_\Sigma \neg G$. On the other hand $G \rhd_\Sigma G$. Ergo $G \rhd_\Sigma \bot$. Also: $\neg G \rhd_\Sigma \Delta \neg G \rhd_\Sigma \neg \Delta G \rhd_\Sigma G$ and $\neg G \rhd_\Sigma \neg G$. So $\neg G \rhd_\Sigma \bot$.

In case T=HA, the results of section 8.1 give us: $(G \vee \neg G) \rhd_\Sigma \bot$. ○

7.6.5 Lindström's Theorem: Per Lindström proved that every equivalence class for interpretability or Π-conservativity contains both a Σ_2- and a Π_2-sentence. We explore his argument in our setting.

We first look at the Σ_2-case. Pick by the Gödel Fixed Point Lemma (GFL) a Σ_2-sentence R with $T \vdash R \leftrightarrow \Box^* B \leq_{ex} \Box^* R$, we have, using the fact that $\rhd_\Sigma$ is a semi-consequence relation for T:

a)	$R \rhd_\Sigma (\Box^* B \leq_{un} \Box^* R \rightarrow B)$	7.6.2
b)	$R \rhd_\Sigma (\Box^* B \leq_{ex} \Box^* R \rightarrow B)$	a, O1, A1, A2
c)	$R \rhd_\Sigma \Box^* B \leq_{ex} \Box^* R$	GFL, A1
d)	$R \rhd_\Sigma ((\Box^* B \leq_{ex} \Box^* R) \wedge (\Box^* B \leq_{ex} \Box^* R \rightarrow B))$	b, c, A3
e)	$R \rhd_\Sigma B$	d, A1, A2
f)	$B \rhd_\Sigma (\neg R \rightarrow \Box^* B \leq_{ex} \Box^* R)$	7.6.1(ii)
g)	$B \rhd_\Sigma \neg\neg R$	GFL, A1, A2

So we find: $R \rhd_\Sigma B \rhd_\Sigma \neg\neg R$. It is open whether we can find Q in Σ_2, such that $B \equiv_\Sigma Q$.

We turn to the $\prod_2$-case. Pick by the Gödel Fixed Point Lemma a $\prod_2$-sentence R with $T\vdash R\leftrightarrow\Box^*B\leq_{un}\Box^*R$, we have:

a)	$R \rhd_\Sigma (\Box^*B\leq_{un}\Box^*R \rightarrow B)$	7.6.2
b)	$R \rhd_\Sigma \Box^*B\leq_{un}\Box^*R$	GFL, A1
c)	$R \rhd_\Sigma ((\Box^*B\leq_{un}\Box^*R)\wedge(\Box^*B\leq_{un}\Box^*R \rightarrow B))$	a, b, A3
d)	$R \rhd_\Sigma B$	c, A1, A2
e)	$B \rhd_\Sigma ((R\rightarrow\Box^*B\leq_{un}\Box^*R)\rightarrow\Box^*B\leq_{un}\Box^*R)$	7.6.1(iv)
f)	$B \rhd_\Sigma R$	GFL, A1, A2
g)	$B \equiv_\Sigma R$	d, f

So every equivalence class contains a $\prod_2$-sentence. ○

7.6.6 Complexity of Σ-preservativity: Let U be the theory axiomatized by the set of A such that $T\vdash A^{\neg\neg}$. Then U is a consistent theory extending PA in the language of PA. $\rhd^*$ stands for: conservativity. Let P_0 be a $\prod_1$-sentence, say P_0 is equivalent to $\neg S_0$, for $S_0\in\Sigma_1$. We have:

$$\begin{aligned} P_0\rhd_{T,\Sigma}\perp &\Leftrightarrow \forall S\in\Sigma_1\,(\,\Box_T(S\rightarrow P_0)\Rightarrow\Box_T\neg S\,)\\ &\Leftrightarrow \forall P\in\prod_1\,(\,\Box_T(S_0\rightarrow P)\Rightarrow\Box_T P\,)\\ &\Leftrightarrow \forall P\in\prod_1\,(\,\Box_U(S_0\rightarrow P)\Rightarrow\Box_U P\,)\\ &\Leftrightarrow T\rhd^*_{U,\prod}S_0. \end{aligned}$$

By a result, due independently to Per Lindström and Robert Solovay, the set $\{S_0\in\Sigma_1 \mid T\rhd^*_{U,\prod}S_0\}$ is complete $\prod_2$. It follows that $\{P_0\in\prod_1 \mid P_0\rhd_{T,\Sigma}\perp\}$ is complete $\prod_2$. ○

8 Closure properties of Σ-preservativity over HA: In this section we verify two closure properties of Σ-preservativity.

8.1 Closure under B1: We will show that Σ-preservativity is closed under B1.

We produce both proofs known to us. The first employs q-realizability. This form of realizability is a translation from $\mathfrak{A}$ to $\mathfrak{A}$, due to Kleene. It is defined as follows:

- $xqP := P$, for P atomic
- $xq(A\wedge B) := ((x)_0qA \wedge (x)_1qB)$
- $xq(A\vee B) := (\,((x)_0=0 \rightarrow (x)_1qA) \wedge ((x)_0\neq 0 \rightarrow (x)_1qB)\,)$
- $xq(A\rightarrow B) := \forall y\,(\,yqA \rightarrow \exists z(\{x\}y\equiv z \wedge zqB)\,) \wedge (A\rightarrow B)$
- $xq\exists yA(y) := (x)_0qA((x)_1)$
- $xq\forall yA(y) := \forall y\,\exists z\,(\{x\}y\equiv z \wedge zqA(y))\,)$

The following facts can be verified in HA:

a) $HA \vdash xqA \rightarrow A$,

b) for every $A \in \Sigma$ and $\{y_1,...y_n\}$ with $FV(A) \subseteq \{y_1,...y_n\}$ we can find an index e such that: $HA \vdash A \rightarrow \exists z\,(\{e\}(y_1,...y_n) \equiv z \wedge zqA)$.

c) Suppose $B_1,...B_n,C$ have free variables among $\{y_1,...y_m\}$ and that $\{x_1,...x_n\}$ is disjoint from $\{y_1,...y_m\}$. Then:
$B_1,...B_n \vdash_{HA} C \Rightarrow \exists e\,[x_1qB_1,..,x_nqB_n \vdash_{HA} \exists z(\{e\}(x_1,...x_n,y_1,...y_m) \equiv z \wedge zqC)]$.

The proofs are all simple inductions. For details the reader is referred to Troelstra[73], 188-202.

We will now show that Σ-preservativity satisfies B1. As is easily seen our proof is verifiable in HA.

We reason as follows:

α)	$A \rhd_{HA,\Sigma} C$	Assumption
β)	$B \rhd_{HA,\Sigma} C$	Assumption
γ)	$S \in \Sigma$ and $HA \vdash S \rightarrow (A \vee B)$	Assumption.
δ)	$xqS \vdash_{HA} \exists z\,(\{e\}(x) \equiv z \wedge zq(A \vee B))$	γ, e provided by (c)
ϵ)	$S \vdash_{HA} \exists z\,(\{e\}(f) \equiv z \wedge zq(A \vee B))$	δ, f provided by (b)
η)	$S \wedge \exists z\,(\{e\}(f) \equiv z \wedge (z)_0 = 0) \vdash_{HA} A$	ϵ
ζ)	$S \wedge \exists z\,(\{e\}(f) \equiv z \wedge (z)_0 \neq 0) \vdash_{HA} B$	ϵ
ϑ)	$S \wedge \exists z\,(\{e\}(f) \equiv z \wedge (z)_0 = 0) \vdash_{HA} C$	η, α
ι)	$S \wedge \exists z\,(\{e\}(f) \equiv z \wedge (z)_0 \neq 0) \vdash_{HA} C$	ζ, β
κ)	$S \vdash_{HA} \exists z\,(\{e\}(f) \equiv z \wedge (z)_0 = 0) \vee \exists z\,(\{e\}(f) \equiv z \wedge (z)_0 \neq 0)$	ϵ
λ)	$S \vdash_{HA} C$	ϑ,ι,κ ❑

The second proof employs a translation due to Dick de Jongh. For later use our definition is slightly more general than is really needed for the problem at hand. Let C be an $\mathfrak{A}$-sentence and let n be a natural number. Define a translation $[C]_n(.)$ as follows:

- $[C]_nP := P$ for P atomic,
- $[C]_n(.)$ commutes $\wedge$, $\vee$, $\exists$,
- $[C]_n(A \rightarrow B) := ([C]_nA \rightarrow [C]_nB) \wedge \Box_n(C \rightarrow (A \rightarrow B))$,
- $[C]_n \forall yA(y) := \forall y\,[C]_nA(y) \wedge \Box_n(C \rightarrow \forall yA(y))$.

Let's first make a few quick observations, that make life easier:

i) $HA \vdash [C]_nA \rightarrow \Box_n(C \rightarrow A)$

ii) $HA \vdash [C]_n((A \rightarrow B) \wedge (A' \rightarrow B')) \leftrightarrow$
$([C]_nA \rightarrow [C]_nB) \wedge ([C]_nA' \rightarrow [C]_nB') \wedge \Box_n(C \rightarrow ((A \rightarrow B) \wedge (A' \rightarrow B')))$.

Similarly for conjunctions of more than two implications.

iii) $HA\vdash [C]_n\forall y\forall z A(y,z) \leftrightarrow (\forall y\forall z [C]_n A(y,z) \wedge \Box_n(C\rightarrow\forall y\forall z A(y,z)))$.
Similary for larger blocks of universal quantifiers.

iv) $HA\vdash [C]_n\forall y(A(y)\rightarrow B(y)) \leftrightarrow$
$\forall y([C]_n A(y) \rightarrow [C]_n B(y)) \wedge \Box_n(C\rightarrow\forall y(A(y)\rightarrow B(y)))$.

v) $HA\vdash [C]_n\forall y<z\ A(y) \leftrightarrow \forall y<z\ [C]_n A(y)$.

vi) For $S\in\Sigma$: $HA\vdash S \leftrightarrow [C]_n S$

(v) is immediate from the well know fact that: $HA\vdash \forall y<z\ \Box_n A(y) \rightarrow \Box_n\forall y<z\ A(y)$.
(vi) is immediate from (v).

Let's write $[C]_n\Gamma := \{[C]_n D \mid D\in\Gamma\}$. We have:

vii) $\Gamma\vdash_{HA,n} A \Rightarrow [C]_n\Gamma \vdash_{HA} [C]_n A$ (verifiably in HA).

Proof of (vii): The proof is by induction on the proof witnessing $\Gamma\vdash_{HA,n} A$. We treat two cases.

$\Gamma=\emptyset$ and A is an induction axiom, say for B(x), of HA_n. Clearly $[C]_n A$ is HA-provably equivalent to:

$[\ ([C]_n B(0) \wedge \forall x([C]_n B(x) \rightarrow [C]_n B(x+1)) \wedge \Box_n(C\rightarrow\forall x(B(x) \rightarrow B(x+1)))) \rightarrow$
$(\forall x [C]_n B(x) \wedge \Box_n(C\rightarrow\forall x B(x)))\ ] \wedge \Box_n(C\rightarrow A)$.

We have:

$HA\vdash \Box_n A$, and hence $HA\vdash \Box_n(C\rightarrow A)$.

So it follows that:

$HA\vdash \Box_n(C\rightarrow\forall x(B(x) \rightarrow B(x+1))) \rightarrow \Box_n(C\rightarrow\forall x B(x))$.

Moreover (as an instance of induction for $[C]_n B(x)$):

$HA\vdash ([C]_n B(0) \wedge \forall x([C]_n B(x) \rightarrow [C]_n B(x+1))) \rightarrow \forall x[C]_n B(x)$.

Combining these we find the promised: $HA\vdash [C]_n A$.

Suppose $A=(D\rightarrow E)$ and the last step in the proof was by:

$\Gamma,D\vdash_{HA,n} E \Rightarrow \Gamma\vdash_{HA,n} D\rightarrow E$.

From $\Gamma,D\vdash_{HA,n} E$, we have by the Induction Hypothesis: $[C]_n\Gamma,[C]_n D\vdash_{HA,n} [C]_n E$ and hence: $[C]_n\Gamma\vdash_{HA,n} [C]_n D\rightarrow[C]_n E$. Moreover for some finite $\Gamma_0\subseteq\Gamma$, we have: $\Gamma_0,D\vdash_{HA,n} E$. Let B be the conjunction of the elements of Γ_0. We find:

$[C]_n\Gamma\vdash_{HA,n} \Box_n(C\rightarrow B)$ and $\vdash_{HA,n} \Box_n(B\rightarrow(D\rightarrow E))$.

Hence: $[C]_n\Gamma\vdash_{HA,n} \Box_n(C\rightarrow(D\rightarrow E))$. We may conclude:

$[C]_n\Gamma\vdash_{HA,n} ([C]_n D\rightarrow[C]_n E) \wedge \Box_n(C\rightarrow(D\rightarrow E))$ ❑(vii)

We now prove our principle. As is easily seen the argument can be verified in HA.

α)	$A \rhd_{HA,\Sigma} C$	Assumption
β)	$B \rhd_{HA,\Sigma} C$	Assumption
γ)	$S \in \Sigma$ and $HA \vdash S \to (A \vee B)$	Assumption.
δ)	for some n $S \vdash_{HA,n} (A \vee B)$	γ
ϵ)	$[\top]_n S \vdash_{HA} ([\top]_n A \vee [\top]_n B)$	δ, (vii)
η)	$S \vdash_{HA} (\Box_n A \vee \Box_n B)$	ϵ, (vi), (i)
ζ)	$HA \vdash S \to C$	α, β, η, 7.1 ❑

8.2 A closure rule for implication: To formulate our next closure rule it is convenient to work in a conservative extension of HA. Let $\mathfrak{A}^+$ be $\mathfrak{A}$ extended with new predicate symbols (including the 0-ary case) for Σ-formulas. Let $\dagger$ be some assignment of Σ-formulas of $\mathfrak{A}$ of the appropriate arities to the new predicate symbols.

Define: $[.]^\circ_n(.)$ and $\{.\}(.): \mathfrak{A}^+ \to \mathfrak{A}$ as follows:

- $\{A\}P := [A]^\circ_n P := P[\dagger]$ for P atomic,
- $\{A\}(.)$ and $[A]^\circ_n(.)$ commute with $\wedge$, $\vee$ and $\exists$,
- $\{A\}(B \to C) := (A \to (B \to C))[\dagger]$, $[A]^\circ_n(B \to C) := \Box_n((A \to (B \to C))[\dagger])$,
- $\{A\}\forall x B(x) := (A \to \forall x B(x))[\dagger]$, $[A]^\circ_n \forall x B(x) := \Box_n((A \to \forall x B(x))[\dagger])$.

We have for B in $\mathfrak{A}^+$ (in the numbering of principles for [.] of 8.1):

viii) $HA \vdash [A[\dagger]]_n(B[\dagger]) \to [A]^\circ_n B \to \{A\}B$,

ix) $[A]^\circ_n B$ is provably equivalent to a Σ-formula.

The proof of (viii) is an easy induction on B. (ix) is trivial.

8.2.1 Theorem: Suppose X is a set of implications in $\mathfrak{A}^+$ and let B be in $\mathfrak{A}^+$. Say, $A := \bigwedge X$ and let $Y := \{C \mid (C \to D) \in X\} \cup \{B\}$. We have, verifiably in HA:

$$(A \to B)[\dagger] \rhd_{HA,\Sigma} \bigvee \{A\}Y.$$

Note that $\Sigma 8$ of section 7 is an immediate consequence of our principle.

Proof: To avoid heavy notations we suppress '$[\dagger]$'. In the context of HA we assume that an $\mathfrak{A}^+$-formula is automatically translated via $\dagger$ to the corresponding $\mathfrak{A}$-formula. Let S be a Σ-sentence (of $\mathfrak{A}$). Suppose $S \vdash_{HA} (A \to B)$. It follows that for some n $S \vdash_{HA,n} (A \to B)$ and hence by (vii) $[A]_n S \vdash_{HA} [A]_n (A \to B)$. By (ii) and (vi) we find:

$$S \vdash_{HA} (\bigwedge \{[A]_n C \to [A]_n D \mid (C \to D) \in X\} \wedge \Box_n (A \to A)) \to [A]_n B,$$

and so:

$$S \vdash_{HA} \bigwedge \{[A]_n C \to [A]_n D \mid (C \to D) \in X\} \to [A]_n B.$$

It follows by (viii) that:

$S \vdash_{HA} \bigwedge\{[A]^\circ_n C \to [A]_n D \mid (C \to D) \in X\} \to [A]^\circ_n B$.

Since HA is a subtheory of PA, we get:

$S \vdash_{PA} \bigwedge\{[A]^\circ_n C \to [A]_n D \mid (C \to D) \in X\} \to [A]^\circ_n B$.

By classical logic, we get: $S \vdash_{PA} \bigvee[A]^\circ_n Y$. Remember that PA is, verifiably in HA, conservative over HA w.r.t $\prod_2$-sentences (see Friedman[77]). Since $(S \to \bigvee[A]^\circ_n Y)$ is $\prod_2$, we get: $S \vdash_{HA} \bigvee[A]^\circ_n Y$. Ergo by (viii): $S \vdash_{HA} \bigvee\{A\}Y$. ❑

Before closing this section we insert some remarks on the proof.

8.2.2 Remarks on the proof

i) The above proof was obtained after analyzing an argument in De Jongh[82].

ii) The step involving conservativity of PA over HA uses the Gödel-Friedman translation. Closer inspection reveals that the argument at hand just requires the Friedman translation. We give a sketch in 8.2.3. It follows that our results generalize to all essentially reflexive RE extensions T of HA that are closed both under the Friedman and the de Jongh translation.

iii) From a sufficiently abstract perspective it would become clear that our present proof is just a variant of the proof of 4.2. $\{G \mid \Box_n(A \to G)\}$ in the present proof corresponds to the grey part of the Kripke model in the picture below 4.2. $[A]_n(.)$ corresponds to the operation of adding the bottom-node in the picture (via Smoryński's operation) to the grey part. The detour via PA corresponds to the fact that our Kripke model argument is essentially classical. The special behaviour of Σ-sentences under translation corresponds with the fact that whether a Σ-sentence is forced or not (in a model of HA) is dependent just on the node under consideration and not on other nodes.

8.2.3 Appendix to 8.2: We show that the double negation translation can be eliminated from the proof of 8.2.1. We pick up the proof from the point where we have proved:

a) $S \vdash_{HA} \bigwedge\{[A]^\circ_n C \to [A]_n D \mid (C \to D) \in X\} \to [A]^\circ_n B$.

Let $H := \bigvee[A]^\circ_n Y$.

The Friedman translation $(E)^H$ of an arithmetical formula E is (modulo some details to avoid variable-clashes) the result of replacing all atomic formulas P in E by $(P \vee H)$. One easily shows:

b) $HA \vdash E \Rightarrow HA \vdash (E)^H$;

c) $HA \vdash H \to (E)^H$;

d) for $S \in \Sigma$: $HA \vdash (S)^H \leftrightarrow (S \vee H)$.

By (a), (b) and (d) we have:

e) $S \vee H \vdash_{HA} \bigwedge\{(([A]^{\circ}{}_{n}C) \vee H) \rightarrow ([A]_{n}D)^{H} \mid (C \rightarrow D) \in X\} \rightarrow (([A]^{\circ}{}_{n}B) \vee H)$.

By (e) and propositional logic we find:

f) $S \vdash_{HA} \bigwedge\{H \rightarrow ([A]_{n}D)^{H} \mid (C \rightarrow D) \in X\} \rightarrow H$.

So by (f) and (c):

g) $S \vdash_{HA} H$.

Hence by (g) and (viii): $S \vdash_{HA} \bigvee\{A\}Y$. ❑

9 On Σ-substitutions: In this section we prove our main results on Σ-substitutions. Given the results we already have, this is rather easy.

9.1 Fact: Let $f \in \Sigma^{\mathfrak{B}}$. We have:

i) $\rhd_{HA,\Sigma,f}$ is a σ-relation.

ii) $\rhd_{HA,\Sigma,\Sigma}$ is a σ-relation.

Proof: (ii) is a direct consequence of (i). We have:

a) Every preservativity relation, and hence $\rhd_{HA,\Sigma,f}$, satisfies A1-3.

b) $\rhd_{HA,\Sigma,f}$ satisfies B1 in virtue of 8.1

c) $\rhd_{HA,\Sigma,f}$ satisfies B2 by 3.5(iv)

d) $\rhd_{HA,\Sigma,f}$ satisfies B3 by 8.2, noting that $\{.\}(.)$ of 8.2 behaves like $[.](.)$ of section 3 on $B(\Sigma)$. ❑

9.2 Theorem

ii) $\rhd_{ROB} = \rhd_{HA,\Sigma,\Sigma} = \rhd_{HA,\top,\Sigma}$.

Proof: By 5.2(iii) $\rhd_{ROB}$ is the minimal σ-relation. So by 9.1(ii): $\rhd_{ROB} \subseteq \rhd_{HA,\Sigma,\Sigma}$. By 3.5(vi): $\rhd_{HA,\Sigma,\Sigma} \subseteq \rhd_{HA,\top,\Sigma}$. We show $\rhd_{HA,\top,\Sigma} \subseteq \rhd_{ROB}$. Suppose *not* $A \rhd_{ROB} B$. Then by 5.2(i): $A^* \nvdash_{IPC} B$. A^* is IPC-provably equivalent to a disjunction of prime NNIL-formulas. It follows that for some such disjunct, say C: $C \nvdash_{IPC} B$. The Heyting algebra $\mathfrak{H}$ axiomatized by C is prime and RE. By the embedding theorem proved in DV[94] (see section 6.2 of this paper) there is an $f \in \Sigma^{\mathfrak{B}}$, such that $HA^* \vdash C[f]$ and $HA^* \nvdash B[f]$. Since $C[f] \in NNIL(\Sigma)$, we have by the $NNIL(\Sigma)$-conservativity of HA* over HA (proved in Visser[82]; see also section 6.2 for the statement of the full result), that $HA \vdash C[f]$. Since HA is a sub-theory of HA*, we have: $HA \nvdash B[f]$. Since $IPC \vdash C \rightarrow A$, it follows that $HA \vdash A[f]$ and $HA \nvdash B[f]$. We may conclude: *not* $A \rhd_{HA,\top,\Sigma} B$. ❑

10 The closed fragment of the provability logic of HA

Let IPC^{up} be the theory in language $\mathfrak{U} := \mathfrak{L}(\{\Box^n\bot \mid 0 \neq n \in \omega\})$ axiomatized by IPC plus:

$\Box^n\bot \vdash \Box^{n+1}\bot$.

We identify $\bot$ with $\Box^0\bot$ and $\top$ with $\Box^\omega\bot$. Let $\mathfrak{V} := \{\Box^n\bot \mid 0 \leq n \leq \omega\}$. The standard interpretation u of $\mathfrak{U}$ in $\mathfrak{A}$ is the interpretation mapping $\Box^n\bot$ to $\Box_{HA}{}^n\bot$.

We will in this section notationally confuse $\Box$ and $\Box_{HA}$. We write $\Box^*$ for $\Box_{HA^*}$.

We reason as follows:

a	for every $A \in NNIL(\Sigma)$ $HA \vdash \Box A \leftrightarrow \Box^*A$	Visser[85]
b	for every $B \in B(\Sigma)$ $HA \vdash \Box B \leftrightarrow \Box B^*$	section 9
c	for every $C \in B(\Sigma)$ $HA \vdash \Box C \leftrightarrow \Box^*C^*$	a,b ($C^* \in NNIL(\Sigma)$)
d	for every $\alpha \in \omega \cup \{\omega\}$ $HA \vdash \Box^\alpha\bot \leftrightarrow \Box^{*\alpha}\bot$	a
e	for every $D \in B(\mathfrak{V})$ $HA \vdash \Box^*(D \leftrightarrow \Box^{\alpha^*(D)}\bot)$	d, section 6.2
f	for every $D \in B(\mathfrak{V})$ $HA \vdash \Box^*D \leftrightarrow \Box^{1+\alpha^*(D)}\bot$	e, a
g	for every $D \in B(\mathfrak{V})$ $HA \vdash \Box D \leftrightarrow \Box^{1+\alpha^*(D^*)}\bot$	f, c
h	for every $E \in \mathfrak{L}_\Box(\emptyset)$ $\exists\beta$ $HA \vdash \Box E \leftrightarrow \Box^\beta\bot$	g

The last step is by induction on the box-depth of E. Note that we can read off from our proof an algorithm to compute β from E.

10.1 Sample computations

i) Consider $A := \neg\Box\bot \rightarrow \Box\Box\bot$. $A^* = \Box\bot \vee \Box\Box\bot$, which is equivalent to $\Box\Box\bot$. Hence:

$HA \vdash \Box A \leftrightarrow \Box^3\bot$.

ii) Consider $A = (\neg\neg\Box\Box\bot \rightarrow \Box\Box\bot) \rightarrow (\Box\bot \vee \neg\Box\bot)$. Let $B := \neg\neg\Box\Box\bot \rightarrow \Box\Box\bot$. We compute A^*.

$$A \equiv ((\Box\Box\bot \rightarrow (\Box\bot \vee \neg\Box\bot)) \wedge ([B](\neg\neg\Box\Box\bot) \vee ([B]\Box\bot) \vee ([B]\neg\Box\bot)))$$
$$\equiv ((\Box\Box\bot \rightarrow (\Box\bot \vee \neg\Box\bot)) \wedge (\neg\neg\Box\Box\bot \vee \Box\bot \vee \neg\Box\bot))$$
$$\equiv ((\Box\Box\bot \rightarrow (\Box\bot \vee \neg\Box\bot)) \wedge (\Box\Box\bot \vee \Box\bot \vee \neg\Box\bot)).$$

Clearly our last formula is equivalent to $\Box\bot \vee \neg\Box\bot$, so via the HA* route we find:

$HA \vdash \Box A \leftrightarrow \Box^2\bot$.

iii) Consider $A = (\neg\neg\Box\bot \rightarrow \Box\bot) \rightarrow (\Box\Box\bot \vee \neg\Box\Box\bot)$. Let $B := \neg\neg\Box\bot \rightarrow \Box\bot$. We compute A^*, using some shortcuts involving IPC^{up}-principles.

$$A \equiv ((\Box\bot \rightarrow (\Box\Box\bot \vee \neg\Box\Box\bot)) \wedge ([B](\neg\neg\Box\bot) \vee ([B]\Box\Box\bot) \vee ([B]\neg\Box\Box\bot)))$$
$$\equiv (\neg\neg\Box\bot \vee \Box\Box\bot \vee \neg\Box\Box\bot)$$
$$\equiv (\Box\bot \vee \Box\Box\bot \vee \neg\Box\Box\bot).$$

Clearly our last formula is equivalent to $\Box\Box\bot \vee \neg\Box\Box\bot$, so via the HA* route we find:

$HA \vdash \Box A \leftrightarrow \Box^3\bot$. ○

Define for $A \in \mathfrak{U}$: $\alpha(A) := \alpha^*(A^*)$. We show that $\alpha(A)$ has a clear meaning in terms of the propositional theory IPC^{up}: it is the maximum of the α such that:

$IPC^{up} \vdash \Box^\alpha \bot \rightarrow A$.

10.2 **Fact:** For $\alpha \in \{0,...,\omega\}$ and $A \in \mathfrak{U}$:

$IPC^{up} \vdash \Box^\alpha \bot \rightarrow A \Leftrightarrow \alpha \leq \alpha(A)$.

Proof: We have:

$IPC^{up} \vdash \Box^\alpha \bot \rightarrow A \Rightarrow B, \Box^\alpha \bot \vdash_{IPC} A$, for some conjunction B of formulas of the form $\Box^\beta \bot \rightarrow \Box^{\beta+1} \bot$.

Since B is in NNIL, we find:

$B, \Box^\alpha \bot \vdash_{IPC} A \Leftrightarrow B, \Box^\alpha \bot \vdash_{IPC} A^*$.

Hence:

$IPC^{up} \vdash \Box^\alpha \bot \rightarrow A \Leftrightarrow IPC^{up} \vdash \Box^\alpha \bot \rightarrow A^*$.

It follows that it is sufficient to prove:

For $\alpha \in \{0,...,\omega\}$ and $A \in NNIL(\{\Box^n \bot \mid 1 \leq n\})$:

$IPC^{up} \vdash \Box^\alpha \bot \rightarrow A \Leftrightarrow \alpha \leq \alpha^*(A)$.

The proof is by induction on A. We leave the simple argument to the reader. The case of disjunction uses the fact that $IPC^{up} + \Box^\alpha \bot$ has the disjunction property. ❑

References

Beeson, M.J., 1975, *The nonderivability in intuitionistic formal systems of theorems on the continuity of effective operations,* JSL 40, 321-346.

Berarducci, A., 1990, *The interpretability logic of Peano arithmetic,* JSL 55, 1059-1089.

Boolos, G., 1993, *The logic of provability,* Cambridge University Press, Cambridge, UK.

De Jongh, D.H.J., 1982, *Formulas of one propositional variable in intuitionistic arithmetic,* in: van TV[82], 51-64.

De Jongh, D.H.J., Renardel, G.R., Van Benthem, J.F.A.K., Visser, A., 1994, *NNIL, a study in intuitionistic propositional logic,* Logic Group Preprint Series 111, Dept. of Philosophy, University of Utrecht, Heidelberglaan 8, 3584 CS Utrecht.

De Jongh, D.H.J., Visser, A., 1994, *Embeddings of Heyting Algebras, revised version,* Logic Group Preprint Series 115, Dept. of Philosophy, University of Utrecht, Heidelberglaan 8, 3584 CS Utrecht.

De Jongh, D.H.J., Chagrova, L.A., to appear, *The decidability of dependency in intuitionistic propositional logic.*

Esakia, L., Artemov, S., 1991, *Provability logic,* Special issue of Studia Logica, L1.

Fagin, R., Halpern, J.Y., Vardi, M.Y., 1992, *What is an inference rule?* JSL, vol 57, 1018-1045.

Friedman, H., 1975, *One hundred and two problems in mathematical logic,* JSL 40, 113-129.

Friedman, H., 1977, *Classically and intuitionistically provably recursive functions,* in: MS[77], 21-27.

Leivant, D., 1975, *Absoluteness in intuitionistic logic,* PhD Thesis, University of Amsterdam. (Corrected reprint: 1979, Mathematical Centre Tract, no 73, Amsterdam.)

Leivant, D., 1980, *Innocuous substitutions,* JSL 45, 363-368.

Leivant, D., 1981, *Implicational complexity in intuitionistic arithmetic,* JSL 46, 240-248.

Müller, G.H., and Scott, D. (eds.), 1977, *Higher set theory,* Springer Lecture Notes in Mathematics 669, Springer, Berlin.

Petkov, P.P. (ed.), 1990, *Mathematical Logic,* (Proceedings of the Heyting Conference, Chaika, 1988), Plenum press, New York and London.

Pitts, A., 1992, *On an interpretation of second order quantification in first order intuitionistic propositional logic,* JSL 57, 33-52.

Renardel de Lavalette, G.R., 1986, *Interpolation in a fragment of intuitionistic propositional logic,* Logic Group Preprint Series 5, Dept. of Philosophy, University of Utrecht, Heidelberglaan 8, 3584 CS Utrecht.

Rybakov, V.V., 1992, *Rules of inference with parameters for intuitionistic logic*, JSL 57, 912-923.

Shavrukov, V., 1993, *Subalgebras of diagonalizable algebras of theories containing arithmetic,* Dissertationes Mathematicae, Polska Akademia Nauk., Mathematical Institute.

Statman, R., 1979, *Intuitionistic propositional logic is polynomial space complete,* Theoretical Computer Science, vol. 9, 67-72.

Švejdar, V., 1983, *Modal Analysis of Generalized Rosser Sentences,* JSL 48, 986-999.

Smoryński, C., 1973, *Applications of Kripke Models*, in Troelstra[73], 324-391.

Smoryński, C., 1985, *Self-Reference and Modal Logic,* Springer Verlag, Berlin

Troelstra, A.S. (ed.), 1973, *Metamathematical Investigations of Intuitionistic Arithmetic and Analysis*, Springer Lecture Notes 344, Springer Verlag, Berlin.

Troelstra, A.S. and Van Dalen, D (eds.), 1982, *The L.E.J. Brouwer Centenary Symposium*, North Holland, Amsterdam.

Troelstra, A.S. and Van Dalen, D, 1988a, *Constructivism in Mathematics*, vol 1, North Holland, Amsterdam.

Troelstra, A.S. and Van Dalen, D, 1988b, *Constructivism in Mathematics*, vol 2, North Holland, Amsterdam.

Van Benthem, J.F.A.K., 1991, *Temporal Logic,* Report X-91-05, ILLC, University of Amsterdam; to appear in *Handbook of Logic in AI and Logic Programming* (Dov Gabbay et al., eds., Oxford University Press).

Van Oosten, J., 1991, *Exercises in realizability,* PhD Thesis, University of Amsterdam.

Visser, A., 1982, *On the Completeness Principle,* Annals of Mathematical Logic 22, 263-295.

Visser, A., 1985, *Evaluation, provably deductive equivalence in Heyting's Arithmetic of substitution instances of propositional formulas,* Logic Group Preprint Series 4, Dept. of Philosophy, University of Utrecht, Heidelberglaan 8, 3584 CS Utrecht.

Visser, A., 1990, *Interpretability Logic,* in:.Petkov[90], 175-209.

Zambella, D., 1994, *Shavrukov's Theorem on the subalgebras of diagonalizable algebras for theories containing $I\Delta_0$+Exp,* Notre Dame Journal of Formal Logic, vol.35, 147-157.