Institute for Language, Logic and Information

ON SOLOVAY'S COMPLETENESS THEOREM

Marc Jumelet

ITLI Prepublications
X-88-01

University of Amsterdam

Institute for Language, Logic and Information
Instituut voor Taal, Logica en Informatie

ON SOLOVAY'S COMPLETENESS THEOREM

Marc Jumelet
Department of Mathematics and Computer Science, Amsterdam

Received November 1988

Master's Thesis
supervisor: D. de Jongh

Correspondence to:

Faculteit der Wiskunde en Informatica
(Department of Mathematics and Computer Science)
Roetersstraat 15
1018WB Amsterdam

or

Faculteit der Wijsbegeerte
(Department of Philosophy)
Grimburgwal 10
1012GA Amsterdam

0. Introduction. This paper has a dual purpose. The first one is the exposition of a modification of a proof of Solovay's first completeness theorem for PA. We will give a method to prove this theorem which does not use the recursion theorem and which clarifies the arithmetical presuppositions underlying the proof. This will be done in chapter 2. Chapter 1 is included merely to introduce the subject matter and to provide two examples which are the starting point of the second subject of this thesis, to be treated in chapter 3. This second subject is a partially successful attempt to strengthen Solovay's arithmetical completeness theorem to infinite sets of formulae.

1. Interpretations of modal logic in arithmetic.

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

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

$P \subseteq SEN_{L\Box}$,

$\bot \in SEN_{L\Box}$,

$\varphi, \psi \in SEN_\Box \Rightarrow (\varphi \rightarrow \psi) \in SEN_{L\Box}$,

and $\varphi \in SEN_{L\Box} \Rightarrow \Box\varphi \in SEN_{L\Box}$.

Boolean connectives $\vee$, $\wedge$, $\neg$, $\leftrightarrow$ will be used as abbreviations with their standard meaning. Instead of $(\Box(\varphi \rightarrow \bot) \rightarrow \bot)$ we will sometimes write $\Diamond\varphi$. It is common practice to let P contain infinitely many symbols. We adopt this convention here, unless it is explicitly stated that we study some finite set of propositional letters.

1.2. The semantics for modal formulae is developed by means of so-called *Kripke-models*. A *model* M *for* $L_\Box$ is a triple $\langle M, R, \Vdash \rangle$, where M is a non-empty set, R a binary relation on M and $\Vdash$ some subset of MxP. $\langle M, R \rangle$ is called the *frame* F of the model. We can uniquely extend the forcing relation to all modal formulae in the following manner (writing $x \Vdash \varphi$ for $\langle x, \varphi \rangle \in \Vdash$ and $x \nVdash \varphi$ for $\langle x, \varphi \rangle \notin \Vdash$):

for all $x \in M$:

for $\chi = p$ for some propositional atom $p \in P$: $x \Vdash p$ iff $x \Vdash p$ in the original sense,
for $\chi = \varphi \to \psi$: $x \Vdash \chi$ iff $x \nVdash \varphi$ or $x \Vdash \psi$,
for $\chi = \Box\varphi$: $x \Vdash \chi$ iff for all $y \in M$ such that xRy: $y \Vdash \varphi$,
and, finally $x \nVdash \bot$.

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

all tautologies of propositional logic;
all expressions of the form $\Box\varphi \to \Box\Box\varphi$,
$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, or $\Box(\Box\varphi \to \varphi) \to \Box\varphi$;
and closed under the following two rules of inference:
$\vdash \varphi \Rightarrow \vdash \Box\varphi$;
if $\vdash \varphi \to \psi$ and $\vdash \varphi$, then $\vdash \psi$.

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

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

(i) M is finite, say $M = \{1, \ldots, n\}$;
(ii) R is a transitive and conversely well-founded relation on M, this means:
$\forall x,y,z \in M (xRy \wedge yRz \to xRz)$ and no infinite ascending chain $x_0Rx_1Rx_2\ldots$ of elements of M exists;
(iii) for all $j \in M$, if $1 < j \leqslant n$, then $1Rj$;
(iv) $1 \Vdash \neg\varphi$.

This theorem is known as the *modal completeness theorem for* **L** with respect to the finite, transitive and conversely well-founded frames. For its proof one may consult Smoryński[1985].
We will now concentrate on models for infinite assumption sets which can consistently be added to the system **L**.

1.5. **Definition.** A set of expressions Δ is called *consistent with respect to* **L** if and only if for no conjunction $\chi_0 \wedge \ldots \wedge \chi_n$ of elements of Δ, $\neg(\chi_0 \wedge \ldots \wedge \chi_n)$ is provable in **L**. We will simply write $\Delta \nvdash_{\mathbf{L}} \bot$ for "Δ is consistent with respect to **L**".

Unfortunately, we cannot hope to prove strong completeness of L. That is, for certain assumption sets Δ such that $\Delta \nvdash_L \bot$, we need models M which necessarily contain infinite ascending R-sequences if we want all formulae of Δ to be forced in some node of the model. These models therefore lack the property of converse well-foundedness.

We will now give two examples of such infinite assumption sets, neither of which can be forced in one single node of a conversely well-founded model.

1.6. Example. Let P, the set of propositional letters, be infinite, say: $P := \{p_0, p_1, p_2, \ldots\}$. Consider the following infinite set of modal expressions (writing $\boxdot\varphi$ for $\varphi \wedge \Box\varphi$):

$\Delta := \{ \boxdot(\Box p_{n+1} \rightarrow p_n) \mid n \in \mathbb{N} \} \cup \{ \neg p_0 \}$.

We claim that this set Δ of modal expressions is consistent with respect to L. Moreover, any model which contains a node in which all expressions occurring in Δ are to be forced, is bound to lack the property of converse well-foundedness.

Suppose $\Delta \vdash_L \bot$. In that case, we would have:

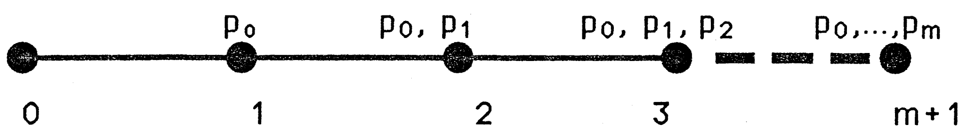
$L \vdash [\boxdot(\Box p_1 \rightarrow p_0) \wedge \ldots \wedge \boxdot(\Box p_{m+1} \rightarrow p_m)] \rightarrow p_0$, for some $m \geqslant 0$. However, we can define a finite model $M = \langle M, R, \Vdash \rangle$ as follows:

$M = \{0, \ldots, m+1\}$,

$\forall x,y \in M\ (xRy \leftrightarrow x < y)$,

$\forall x \in M \forall n \in \mathbb{N}\ (x \Vdash p_n \leftrightarrow x > n)$.

It is clear now, that $0 \Vdash \boxdot(\Box p_1 \rightarrow p_0) \wedge \ldots \wedge \boxdot(\Box p_{m+1} \rightarrow p_m)$ holds, whereas, on the other hand, $0 \Vdash \neg p_0$ is the case. By theorem 1.4 we can conclude that our assumption is absurd. It follows, that Δ is consistent. The diagram summarizes the proof (atoms shown only when forced):



Nevertheless, no transitive and conversely well-founded frame can provide a model for all formulae of Δ if they are all to be forced in some node of it. To see this, we assume the existence of such a model $M := \langle M, R, \Vdash \rangle$, R transitive, which contains a node x_0

such that $x_0 \Vdash \varphi$ for all $\varphi \in \Delta$, and construct an infinite R-sequence as follows:

$x_0 \Vdash \neg p_0$, since $\neg p_0 \in \Delta$.

Suppose x_i to be found such that $x_i \Vdash \neg p_i$ and $x_i = x_0$ or $x_0 R x_i$. In that case some $x_{i+1} \in M$, satisfying $x_{i+1} \Vdash \neg p_{i+1}$ and $x_i R x_{i+1}$ is bound to exist, since $x_i \Vdash \neg\Box p_{i+1}$ is a direct consequence of the fact that the formula $\boxdot(\Box p_{i+1} \to p_i)$ occurs in Δ and is therefore forced in x_0. But now we have $x_0 R x_{i+1}$ by transitivity, so we can repeat the process.

This construction produces an infinite R-sequence $x_0 R x_1 R x_2 \ldots$ of elements of M.

The next example was suggested by de Jongh. It shows us that we can obtain a similar negative result, using only a finite stock of propositional letters.

1.7. Example. We can define a sequence $\langle\varphi_n\rangle_n$ of modal expressions containing but a single propositional letter p as follows:

$\varphi_0 := p \wedge \Box\bot$;

$\varphi_n := p \wedge \Box\neg p \wedge \neg\Box^n\bot \wedge \Box^{n+1}\bot$ for $n > 0$;

where $\Box^n\chi$ denotes the formula χ prefixed by n boxes.

With this sequence we define an infinite assumption set Δ which has the same properties as the one we used in the example above:

$\Delta := \{\boxdot(\Diamond\varphi_i \to \Diamond\Diamond\varphi_{i+1}) \mid i \in \mathbb{N}\} \cup \{\Diamond\Diamond\varphi_0\}$.

Statement: $\Delta \nvdash_L \bot$. Suppose we could derive a contradiction from Δ in L. Now let $\{\boxdot(\Diamond\varphi_i \to \Diamond\Diamond\varphi_{i+1}) \mid i \leqslant m\} \cup \{\Diamond\Diamond\varphi_0\}$ for some $m \in \mathbb{N}$ be the finite subset of Δ responsible for this contradiction. By theorem 1.4 we are done, once we have constructed a finite, transitive and conversely well-founded model $\mathbf{M} = \langle M, R, \Vdash\rangle$ which verifies this finite subset. We define:

$M := \{\langle x,y\rangle \mid x,y \in \mathbb{N} \wedge y \leqslant x \wedge x \leqslant m+2\}$. It goes without saying that we could have used any numbering of the nodes of the model.

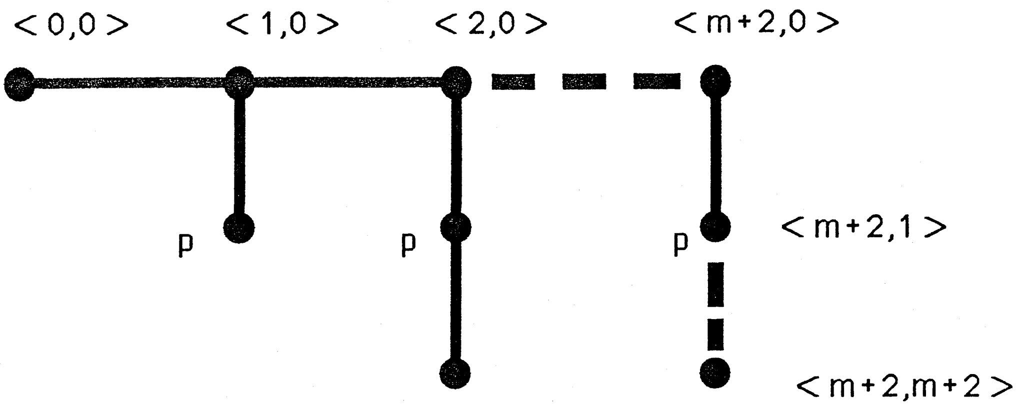
To define R on M we set:

$\langle x,y\rangle R \langle x',y'\rangle \leftrightarrow (y=0 \wedge x < x') \vee (x=x' \wedge y < y')$.

It is readily observed, that R is transitive and well-founded on M. For the forcing relation we define:

$\langle x,y\rangle \Vdash p$ iff $y=1$.

The situation thus obtained may be displayed as in the diagram below:



From the definition of the various φ_n's we can conclude: $\langle x,y\rangle \Vdash \varphi_i$ if and only if $y=1 \wedge x=i+1$. So $\langle 0,0\rangle \Vdash \Diamond\varphi_i$ for all i such that $0 \leqslant i \leqslant m+1$. But also $\langle 0,0\rangle \Vdash \Diamond\Diamond\varphi_{i+1}$ for all i such that $0 \leqslant i \leqslant m+1$ holds, so we may safely conclude that, for any relevant i, $\langle 0,0\rangle \Vdash \Diamond\varphi_i \rightarrow \Diamond\Diamond\varphi_{i+1}$. The same type of argument applies to the axioms of the form $\Box(\Diamond\varphi_i \rightarrow \Diamond\Diamond\varphi_{i+1})$. Suppose $\langle x,y\rangle \Vdash \Diamond\varphi_i$ for $\langle x,y\rangle \neq \langle 0,0\rangle$. In that case we would have $\langle x',y'\rangle \Vdash \varphi_i$ for some $\langle x',y'\rangle$ satisfying $\langle x,y\rangle R \langle x',y'\rangle$, hence $\langle x,y\rangle R \langle i+1,1\rangle$. But then we have: $y=0 \wedge x \leqslant i+1$, from which we can conclude $\langle x,y\rangle R \langle i+2,0\rangle$. Since $\langle i+2,0\rangle \Vdash \Diamond\varphi_{i+1}$ will hold, we can conclude: $\langle x,y\rangle \Vdash \Diamond\Diamond\varphi_{i+1}$. On the other hand, we also have $\langle 0,0\rangle \Vdash \Diamond\Diamond\varphi_0$, thus the model verifies the given initial segment of Δ. Therefore, Δ is consistent with respect to L. But, if we are to construct a model on which all formulae of Δ are verified, we run into the same difficulty as in the previous example. Suppose all formulae of Δ to be true in some node x_0 of a model. Apparently, we would have $x_0 \Vdash \Diamond\Diamond\varphi_0$. But then, because $\Box(\Diamond\varphi_0 \rightarrow \Diamond\Diamond\varphi_1)$ is in Δ, there should be some x_1 in the model such that $x_0 R x_1$ and $x_1 \Vdash \Diamond\varphi_0 \wedge \Diamond\Diamond\varphi_1$. Again, there should be some x_2 in the model such that $x_1 R x_2$ and $x_2 \Vdash \Diamond\varphi_1 \wedge \Diamond\Diamond\varphi_2$, and so on. In other words, we can construct an infinite R-sequence of elements of this model in just the same fashion as we did in the previous example.

1.8.Interpretations. An interpretation of a set of modal formulae is a function $(\ )^*$ that assigns a sentence φ^* of Peano arithmetic to each modal expression φ and obeys the following criteria:

$(\bot)^* = 0=1;$

$(\varphi \to \psi)^* = \varphi^* \to \psi^*;$

$(\Box\varphi)^* = \exists p \text{ proof } (p, \ulcorner \varphi^* \urcorner).$

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

In chapter 3 the substitutionary nature of the interpretation function $(\ \ )^*$ will become a matter of interest. As it is often used implicitly, the next trivial lemma is formulated.

1.9.Lemma. For every set of modal propositional variables P and every interpretation function $(\ \ )^*$, the following holds:

Let, for each p in P, a sentence s_p of arithmetic be given, such that $PA \vdash s_p \leftrightarrow p^*$. Then, for every modal expression φ, the formula $\varphi^* \leftrightarrow \varphi^{**}$ is provable in Peano, if $(\ )^{**}$ is defined by: $p^{**} = s_p$.

The proof of this lemma is by induction on the length of the modal formulae.

1.10. Solovay's first Completeness Theorem (Solovay[76]). This theorem is formulated as follows:

Let φ be any modal expression, then: $\vdash_L \varphi$ if and only if $PA \vdash \varphi^*$ for every interpretation $(\ \ )^*$ of the modal language used, provided it satisfies the three clauses of the preceding paragraph.

Naturally there is no need to bother about the number of propositional variables here, because any modal expression can contain only a finite number of them. The implication from the left to the right is of no concern to us here. The proof is simple, due to the fact that Peano arithmetic is closed under the axioms and rules of L whenever the provability predicate is substituted for the modal operator $\Box$. The arithmetical versions of the rules and axioms of L are exactly the three Löb conditions and Löb's theorem which are fulfilled in Peano arithmetic. The conditions imposed upon the interpretation function will do the rest. The

remaining implication will be treated below. The modification of Solovay's proof which we will present below, is based on an idea of Franco Montagna and was further simplified by Dick de Jongh.
In section three we will use Solovay's second completeness theorem:

1.11. Second Completeness Theorem (Solovay[76]).
Let S be the smallest set of modal formulae containing all theorems of L, all formulae of the form $\Box\varphi \rightarrow \varphi$ and closed under modus ponens, then, for all modal formulae φ:

$\varphi \in S$ if and only if $\mathbb{N} \models \varphi^*$ for all $(\ )^*$ satisfying the criteria of paragraph 1.8 (cf. Solovay[76]).

2. A revised proof of Solovay's theorem.

2.1 The proof of the completeness theorem is based on the idea that a certain class of Kripke-models can be embedded in Peano arithmetic. We have already seen that any modal expression φ which is not derivable from the axioms of L, gives rise to the construction of some countermodel on which φ is falsified. The embedding of such a model in its turn was carried out by Solovay by defining, with the aid of the recursion theorem, a recursive function h which paces through the model in a highly peculiar way. Intuitively speaking, one can describe the Solovay function as follows: as values it takes only numbers denoting the nodes of the Kripke-model in question. The next value can only be the same as the previous one or one which is accessible from it by means of the relation R in the model. Thus it is clear that this function eventually reaches a limit. This limit is used to specify more exactly the next value each time, namely in the following way: for each argument the function takes the same value m as the previous one, unless the argument codes a proof in PA of the fact that for a certain number n, R-accessible from m, the limit of the function is not equal to n. It is therefore clear that it is mainly the eventual value of that function, its limit, so to speak, which plays a role.
As the technical part of the proof involves only the mutual relations between these limit assertions, we may be tempted to define corresponding sentences, using nothing but the desired connection with the other sentences. More precisely, we may replace

each expression "$l=i$" we come across in the original proof (the eventual value of h is i), by a single sentence λ_i, the definition of which is an exact imitation of the conditions under which $l=i$ became true. It is important to notice that these conditions can all be spelled out in the form of finite conjunctions, claiming the existence or non-existence and order of succession of certain proofs, namely proofs of other expressions of the form $\neg l=j$. But within proof predicates only codes of these expressions occur. It seems plausible therefore to define each λ_i by means of a fixed-point equation, containing only codes of these λ_j's. It will be demonstrated below, that, in doing so, the alternative sentences satisfy the same lemmas as the original ones did. This makes them equally suitable to perform as arithmetical interpretations of the modal logic.

2.2. Definitions. Let $F=\langle M,R\rangle$ be a finite, transitive and conversely well-founded frame. $M=\{1,\dots,n\}$ and for all j, if $1<j\leqslant n$, then $1Rj$.

We'll use the following abbreviations:

$i\underline{R}j$ for $i=j\vee iRj$;

$i\circ j$ for $\neg i\underline{R}j\wedge\neg j\underline{R}i$.

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

$$PA\vdash\lambda_1\leftrightarrow\Box\neg\lambda_1\wedge\bigwedge_{1Ri}\neg\Box\neg\lambda_j;$$

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

$$PA\vdash\lambda_i\leftrightarrow\Box\neg\lambda_i\wedge\bigwedge_{iRj}\neg\Box\neg\lambda_j\wedge\bigwedge_{i\circ j}\bigvee_{\substack{k\underline{R}i\\ k\circ j}}(\Box\neg\lambda_k\prec\Box\neg\lambda_j).$$

"$\Box A\prec\Box B$" here is the usual notation for:

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

Finally, we define:

$$\lambda_0:=\neg\bigvee_{1\leqslant i\leqslant n}\lambda_i.$$

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

1) $PA\vdash\bigvee_{0\leqslant i\leqslant n}\lambda_i$.

2) $\mathbb{N}\models\lambda_0$.

3) For all i such that $0 \leqslant i \leqslant n$, $PA+\lambda_i$ is consistent.

4) $PA \vdash \lambda_i \rightarrow \bigwedge\!\!\!\bigwedge_{iRj} \neg\Box\neg\lambda_j$ for all $i \geqslant 0$.

5) $PA \vdash \lambda_i \rightarrow \bigwedge\!\!\!\bigwedge_{\neg iRj} \Box\neg\lambda_j$ for all $i > 0$.

This lemma is the main clue to the proof of Solovay's completeness theorem. If we replace each expression of the form λ_i by $l=i$, we get the original lemma (cf. Solovay[76], lemma 4.1).

2.4. Smoothing the proof. For reasons of economy, it is useful to prove lemma 2.3 within a more general framework. This will show us exactly which properties of **PA** are used to prove lemma 2.3. We take for this purpose a modified version of R^-, the modal system of Guaspari and Solovay that accounts for the behaviour of witness-comparison formulae (cf. Guaspari and Solovay[79]).
We first recall that R^- is an extension of **L** in which the class of well-formed formulae is enlarged by the so-called witness-comparison formulae, viz. those of the forms $\Box A \prec \Box B$ and $\Box A \preccurlyeq \Box B$. R^- is axiomatized by adding to **L** the axiom schemata (cf. de Jongh[87]):

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

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

(01) $A \rightarrow A \preccurlyeq B \vee B \preccurlyeq A$;

(02) $A \preccurlyeq B \rightarrow A$;

(03) $A \preccurlyeq B \wedge B \preccurlyeq C \rightarrow A \preccurlyeq C$;

(04) $A \prec B \leftrightarrow A \preccurlyeq B \wedge \neg B \preccurlyeq A$.

We extend R^- as follows: for any $F=\langle M,R \rangle$, being a finite, transitive and conversely well-founded frame, with $M=\{1,\ldots n\}$ and $1Ri$ for all i such that $1 < i \leqslant n$, let R^-_F be defined by adding the following axioms to R^- (we assume the language to contain propositional constants $L_0,\ldots, L_n$):

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

for each i such that $1 < i \leqslant n$:
$\boxdot(L_i \leftrightarrow \Box\neg L_i \wedge \bigwedge_{iRj} \neg\Box\neg L_j \wedge \bigwedge_{i \circ j} \bigvee_{\substack{kBi \\ k \circ j}} (\Box\neg L_k \prec \Box\neg L_j));$

$\boxdot(L_0 \leftrightarrow \neg \bigvee_{1 \leqslant i \leqslant n} L_i).$

These axioms will be referred to as the *limit axioms*. In addition to these, we let R_F^- contain $\boxdot(\neg(\Box\neg L_i \preccurlyeq \Box\neg L_j \wedge \Box\neg L_j \preccurlyeq \Box\neg L_i))$ for all i,j such that $0 \leqslant i,j \leqslant n$ and $i \neq j$, as so-called *proof apartness axioms*. In the next two paragraphs we will mention some properties of R_F^- that will be needed for the proof of lemma 2.3.

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

2.5.Theorem (*Soundness* of R_F^-). An interpretation $(\ )^+$ of sentences in the language of R_F^- into the language of arithmetic is called F-sound if and only if $(\ )^+$ fulfils the criteria cited for $(\ )^*$ in paragraph 1.8 and, in addition to these:
for all formulae φ, ψ:
$(\Box\varphi \preccurlyeq \Box\psi)^+ = \exists p\,[\mathrm{proof}(p, \ulcorner\varphi^+\urcorner) \wedge \neg\exists q < p\ \mathrm{proof}(q, \ulcorner\psi^+\urcorner)];$
$(\Box\varphi \prec \Box\psi)^+ = \exists p\,[\mathrm{proof}(p, \ulcorner\varphi^+\urcorner) \wedge \neg\exists q \leqslant p\ \mathrm{proof}(q, \ulcorner\psi^+\urcorner)];$
for all i such that $0 \leqslant i \leqslant n$:
$L_i^+ = \lambda_i$ (in the sense of definition 2.2.);
Soundness of R_F^- is formulated as follows: for all interpretations $(\ )^+$ of sentences in the language of R_F^- the following holds for any φ in that language: $R_F^- \vdash \varphi \Rightarrow PA \vdash \varphi^+$.

The proof is straightforward by induction on the length of proof in R_F^-, since PA is closed under the same rules and axioms we have at our disposal in R_F^- provided $(\ )^+$ is F-sound. We will use this theorem extensively in the proof of lemma 2.3.

A Kripke-model for R^- is a finite, tree-ordered Kripke-model for L in which witness-comparison formulae are treated as if they were atomic formulae and in which the following two requirements are fulfilled:
if $i \Vdash A \preccurlyeq B$ and iRj, then $j \Vdash A \preccurlyeq B$;
each instance of the order-axioms is fulfilled at each node.

Completeness of R^- is stated as follows:

$R^- \vdash \varphi$ iff φ is valid on all finite, tree-ordered Kripke-models for R^- .

In the case of R_F^-, defined as in paragraph 2.4, this theorem implies:

2.6. Theorem (*completeness* of R_F^-).

If $R_F^- \nvdash \varphi$, then a finite, tree-ordered Kripke-model for R^- exists, in which all limit-axioms and proof-apartness axioms are forced at each node, and on which φ is falsified.

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

$R_F^- \vdash \varphi \Leftrightarrow R^- \vdash \theta \rightarrow \varphi$, where θ is the finite conjunction of limit axioms and proof apartness axioms listed in the definition of R_F^-.

The implication from the right to the left is easily proved. The other direction is proved by induction on the length of proof in R_F^-. To obtain the desired result, we should check whether any proof of a formula φ in R_F^- can be transformed into a proof of $\theta \rightarrow \varphi$ in R^-. But this can cause no difficulty, since any axiom of R_F^- is either an axiom of R^- or a consequence of θ. Besides, if the last rule applied in a proof in R_F^- of some formula ψ had been the necessitation rule (from $\vdash \chi$ infer $\vdash \Box \chi$), then we could use $\theta \rightarrow \Box \theta$ which is a theorem of R^-. ⊠

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

Now we are ready to commence the proof of lemma 2.3.

Proof of lemma 2.3. Fix a finite, transitive and conversely well-founded frame $F=\langle M,R\rangle$, with $M=\{1,...n\}$ and $1Ri$ for all i such that $1 < i \leqslant n$. Let $\lambda_0,...,\lambda_n$ and R_F^- be as in definitions 2.2 and 2.4. We first show:

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

As the implication from the right to the left is obviously provable, we will concentrate on the opposite direction. Suppose the contrary to be the case. By theorem 2.6 we would have a finite, tree-ordered Kripke-model for R^- with limit axioms and proof apartness axioms forced everywhere in the model and with some bottom-node k_0 such that $k_0 \Vdash L_0 \wedge \bigvee_{1 \leqslant i \leqslant n} \Box\neg L_i$.

Now, we must have: $k_0 \Vdash \Box\neg L_{i_1} \wedge \ldots \wedge \Box\neg L_{i_k} \wedge \bigwedge_{j \notin \{i_1 \ldots i_k\}} \neg\Box\neg L_j$, for some k such that $1 \leqslant k \leqslant n$.

First we remark that we are free to replace the symbol "$\prec$" in the limit axioms by "$\preccurlyeq$", due to the fact, that the apartness axioms are forced in k_0.

If $k=1$, we derive a contradiction straightaway, since then we would have $k_0 \Vdash \Box\neg L_{i_1} \wedge \bigwedge_{j \neq i_1} \neg\Box\neg L_j$, which implies a fortiori: $k_0 \Vdash \Box\neg L_i \wedge \bigwedge_{jRi_1} \neg\Box\neg L_j \wedge \bigwedge_{j \circ i_1} \Box\neg L_{i_1} \prec \Box\neg L_j$. But now we obtain: $k_0 \Vdash L_{i_1}$, contradicting $k_0 \Vdash L_0$. We may therefore assume $k > 1$. As any instance of the order axioms is forced at k_0, we can stipulate, without loss of generality, that at k_0 the following is forced:

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

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

$m_1 := 1$;

$m_{h+1} := m$ for m being the smallest index number in $\{1, \ldots k\}$ such $i_{m_h} R\, i_m$ and $k_0 \Vdash \Box\neg L_{i_{m_h}} \preccurlyeq \Box\neg L_{i_m}$. If no such m exists, set $l=h$ and $m_{h+1} = m_h$.

It will be understood that this construction comes to an end in any case, because the set $\{1, \ldots, k\}$ is finite. Again if $l=1$, we obtain an absurd situation. In that case, we would have $k_0 \Vdash \bigwedge_{i_1 R j} \neg\Box\neg L_j$, since otherwise some $j \neq i_1$ had been in $\{i_{m_1}, \ldots, i_{m_l}\}$. Now we have $k_0 \Vdash \bigwedge_{j \circ i_1} \Box\neg L_{i_1} \preccurlyeq \Box\neg L_j$, since L_{i_1} was first in line anyhow. But this immediately leads to $k_0 \Vdash L_{i_1}$, contradicting $k_0 \Vdash L_0$. So we may assume $l > 1$. By means of a finite induction procedure we will now prove the following: for all p such that $1 \leqslant p \leqslant l$:

$$k_0 \Vdash \bigwedge_{j \circ m_p} \bigvee_{\substack{k \circ j \\ k \underline{R} m_p}} (\Box\neg L_k \prec \Box\neg L_j)).$$

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

Induction step: suppose $k_0 \Vdash \bigwedge_{j \circ m_p} \bigvee_{\substack{k \circ j \\ k \underline{R} m_p}} (\Box\neg L_k \prec \Box\neg L_j))$. Now let j be such, that $j \circ m_{p+1}$.

There are two possibilities: either $j \circ m_p$ as well, or not.

In the first case we obtain $k_0 \Vdash \bigvee_{\substack{k \circ j \\ k \underline{R} m_{p+1}}} (\Box\neg L_k \prec \Box\neg L_j))$ by

induction hypothesis, for $k \underline{R} m_p$ implies $k \underline{R} m_{p+1}$.

In the latter case m_pRj must hold. But the definition of m_{p+1} implies: $k_0 \Vdash \Box\neg L_{i_{m_{p+1}}} \preccurlyeq \Box\neg L_j$ whence $k_0 \Vdash \bigvee\limits_{\substack{k \circ j \\ k B m_{p+1}}} (\Box\neg L_k \prec \Box\neg L_j)$ follows by propositional logic.

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

$$k_0 \Vdash \Box\neg L_{i_{m_1}} \wedge \bigwedge\limits_{i_{m_1} R j} \neg\Box\neg L_j \wedge \bigwedge\limits_{j \circ i_{m_1}} \bigvee\limits_{\substack{k \circ j \\ i_{m_1} B j}} (\Box\neg L_k \prec \Box\neg L_j) .$$

But this implies $k_0 \Vdash L_{i_{m_1}}$ contradicting $k_0 \Vdash L_0$. The proof is hereby completed, since nothing specific about the set $\{i_1,\ldots i_k\}$ had been presupposed apart from its being non-empty.

b) If $1 \leqslant i \leqslant n$, then $R_F^- \vdash L_i \rightarrow \bigwedge\limits_{iRj} \neg\Box\neg L_j$. This is immediate from the definition of R_F^- .

Combining a) and b) we get 4) of lemma 2.3 by soundness.

c) R_F^- contains all tautologies of propositional logic, so we have $R_F^- \vdash L_0 \vee \neg L_0$ from which $R_F^- \vdash \bigvee\limits_{0 \leqslant i \leqslant n} L_i$ is readily deduced. Employing soundness, this accounts for 1) of lemma 2.3.

As all theorems of PA hold in the standard model, we must have $\mathbb{N} \models \lambda_i$ for some i such that $0 \leqslant i \leqslant n$. But then $\mathbb{N} \models \lambda_0$ must hold, since for any $i \neq 0$ we would have $PA \vdash \neg\lambda_i$ in case λ_i were true. Combining this with 4) of lemma 2.3, we obtain $\mathbb{N} \models \bigwedge\limits_{0 \leqslant i \leqslant n} \neg\Box\neg\lambda_i$. This settles 2) and 3) of lemma 2.3.

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

By a) we have $R_F^- \vdash \Box\neg L_i \rightarrow \neg L_0$. Applying the necessitation rule $(\vdash\varphi \Rightarrow \vdash\Box\varphi)$ we infer: $R_F^- \vdash \Box\Box\neg L_i \rightarrow \Box\neg L_0$. As $\Box\neg L_i$ is a boxed formula, $\Box\neg L_i \rightarrow \Box\Box\neg L_i$ is a theorem of R_F^-. The proof is now completed, since $R_F^- \vdash L_i \rightarrow \Box\neg L_i$ is a direct consequence of the definition of R_F^-.

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

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

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

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

$$R_F^- \vdash L_i \rightarrow \bigwedge\limits_{i \circ j'} \bigvee\limits_{\substack{k B i \\ k \circ j'}} (\Box\neg L_k \prec \Box\neg L_{j'}).$$

More specifically, we obtain:

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

As the order axioms will arrange the various expressions $\Box \neg L_k$ of the consequent in one way or another, we have:

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

But the consequent in the last formula is a Σ-expression implying $\neg L_j$, so: $R_F^- \vdash \bigwedge_{\substack{i \circ j' \\ j' \underline{B} j}} \bigvee_{\substack{k \underline{B} i \\ k \circ j'}} (\Box \neg L_k \prec \Box \neg L_{j'}) \rightarrow \Box \neg L_j$, which completes this proof.

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

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

Applying soundness, this settles 5) of lemma 2.3. ⊠

2.7. About the completeness theorem. Let $M = \langle M, R, \Vdash \rangle$ be a finite, transitive and conversely well-founded model with $M = \{1, \ldots, n\}$ and for all i if $1 < i \leqslant n$, then 1Ri. As usual, we expand M by adding an extra node 0 to it and defining $0 \Vdash$ as equivalent to $1 \Vdash$ for all propositional letters. By definition 2.2 we obtain sentences $\lambda_0, \ldots, \lambda_n$ satisfying lemma 2.3. We define an interpretation $(\ )^*$ by setting for all $p \in P$:

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

The following lemma provides the necessary clue to the completeness theorem:

Lemma: for all modal expressions φ, if $1 \leqslant i \leqslant n$, then

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

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

The proof is exactly the same as the original one, with each expression of the form $l = i$ replaced by λ_i, so we will not give it here. It will be understood, that in fact any set of sentences $\lambda_0, \ldots, \lambda_n$ of Peano which satisfy the requirements of definition 2.2 can be used to obtain a suitable interpretation.

Our explanation concerning the adapted proof of Solovay's result is now completed.

3. Solovay's theorem for infinite sets of formulae.

3.1. Consistent interpretations of sets of formulae. One of the first questions which may arise within the context of Solovay's completeness theorem with respect to sets of modal formulae, is the following: if Δ is a set of modal formulae such that $\Delta \nvdash_L \bot$, can we define a consistent interpretation (in the sense of paragraph 1.8), such that $\Delta^* = \{\varphi^* | \varphi \in \Delta\}$ is consistent with respect to Peano arithmetic? The answer is simply yes. As a matter of fact, the so-called *uniformisation* of Solovay's completeness theorem gives an interpretation ( )* such that for all modal formulae φ^*, the following holds:

$\vdash_L \varphi$ if and only if $PA \vdash \varphi^*$ (cf. Visser[81] and Artyomov[80]).

This means that, for any consistent set Δ of formulae, the interpretation ( )* gives a consistent set of sentences Δ^*. Actually, e.g. in the example of paragraph 1.6, we would like more, namely an interpretation ( )* which interprets formulae $\Box A$ in Δ not so much as formulae which can be consistently assumed to be provable, but which actually are provable. Similar considerations apply to the second example. A. Visser[88] succeeded in giving an interpretation with the desired properties for the first example. We will sketch here a general method which applies to very well behaved sets of sentences consistent with respect to S (Solovay's extension of L which is arithmetically complete for the formulae which under any interpretation become true sentences). In particular, the method applies to both examples 1.6 and 1.7. Unfortunately, we have as yet been unable to give some nice sufficient conditions for our method to be applicable.

3.2.Example. A unary predicate A(v) exists, satisfying:

$PA \vdash \forall x[\ A(x) \leftrightarrow (\Box A(x+1) \wedge \neg \exists y \leqslant x \text{ proof}(y, \ulcorner A(0) \urcorner))]$.

Applying this predicate, we are able to translate the infinite assumption set Δ of paragraph 1.6. The translated set has already been studied within the context of descending hierarchies of reflection principles (cf. Visser[88]).

Define, for each $i \in \mathbb{N}$: $p_i^* := A(\underline{i})$, where $\underline{i}$ is the numeral corresponding with i.

Claim: $PA \nvdash p_0^*$. Proof. Suppose the contrary to be the case and let q+1 be the code of the shortest proof of A(0). Since for any $q' \leqslant q$

we have: $PA\vdash A(q') \Rightarrow PA\vdash A(q'+1)$ we are forced to conclude: $PA\vdash A(q+1)$. This will immediately lead to:
$PA\vdash \neg proof(q+1,\ulcorner A(0)\urcorner)$, which is absurd. But now we derive: $\forall i\in\mathbb{N}\ PA\vdash \neg\exists y\leqslant i\ proof(y,\ulcorner A(0)\urcorner)$, whence follows: $\forall i\in\mathbb{N}$ $PA\vdash \boxdot(\Box A(i+1)\rightarrow A(i))$. Apparently, the addition of the set of axioms $\{\boxdot(\Box p_{i+1}{}^* \rightarrow p_i{}^*) \mid i\in\mathbb{N}\}$ to Peano is redundant. This insures that the defined interpretation of Δ is a suitable one. ⊠

In order to generalize the above result, we will first concentrate on sequences of models which are to be used in arithmetical interpretations of infinite assumption sets. It is clear that, for any Δ consistent with respect to L, we will have a set of finite models which verify a given initial segment of Δ. But what we need is a sequence $\langle M_n\rangle_n$ of models such that any model in it is an extension of its predecessor and that any finite subset of Δ is verified by some model M_n in the sequence and subsequently by all of its successors. Hence the following definition:

3.3. Definition. $\langle M_n\rangle_n = \langle\langle M_n, R_n, \Vdash_{M_n}\rangle\rangle_n$ is called a *sequence of models for an assumption set* $\Delta=\{\chi_0,\chi_1,\chi_2,\ldots\}$ consistent with respect to L if and only if the following clauses are fulfilled:

for all $n\in\mathbb{N}$:

(a) $M_n\subset M_{n+1}$;

(b) $\forall i\in M_n$: $i\Vdash_{M_n} p \Leftrightarrow i\Vdash_{M_{n+1}} p$;

(c) $\forall i,j\in M_n$: $iR_nj \Leftrightarrow iR_{n+1}j$;

(d) M_n is finite, say $M_n=\{1,\ldots k_n\}$;

(e) R_n is transitive and conversely well-founded on M_n ;

(f) $\forall i\in M_n(i\neq 1 \Rightarrow 1R_ni)$;

(g) $1\Vdash_{M_n} \chi_0\wedge\ldots\wedge\chi_n$.

The following lemma is a kind of strong completeness theorem for L with respect to sequences of models:

3.4.Lemma. Let $\Delta=\{\chi_i \mid i\in\mathbb{N}\}$ be properly infinite, so $\Delta\vdash_L \neg\Box^n\bot$ for all $n\geqslant 1$. If $\Delta\nvdash_L\bot$, then a sequence of models for Δ exists.
Proof. Let in the following Φ_X for any set of modal formulae X, be defined as the set of all subformulae of formulae in X, closed under negation in the following sense:

if $\varphi\in\Phi_X$ and not $\varphi=\neg\psi$ for some formula ψ, then $\neg\varphi\in\Phi_X$.

Assume $\Delta \nvdash_L \bot$. Fix an enumeration of all formulae occurring boxed in Φ_Δ. Let Γ be a maximal L-consistent extension of Δ within Φ_Δ. Set $\Gamma_{\langle\rangle}:=\Gamma$. If Γ_m has been constructed for a finite sequence denoted by m, then we construct $\Gamma_{m*\langle i\rangle}$ from it only if $\neg\Box\varphi_i \notin \Gamma_m$, by taking a maximal consistent extension of the set $\{\Box\psi, \psi \mid \Box\psi \in \Gamma_m\} \cup \{\Box\varphi_i, \neg\varphi_i\}$. Before we define a sequence of models, we will first regroup the maximal consistent extensions. Let for all $n \in \mathbb{N}$, W_n be the smallest set such that:

$\Gamma_{\langle\rangle} \in W_n$;

$\Gamma_{m*\langle i\rangle} \in W_n$ if $\Gamma_m \in W_n$ and $\neg\Box\varphi_i \in \Phi_{\{\chi_0 \dots \chi_n\}}$.

In the first place it is clear, that, for all $n \in \mathbb{N}$, W_n is finite, since the number of formulae of the form $\neg\Box\varphi$ which are in $\Phi_{\{\chi_0 \dots \chi_n\}}$ is eventually exhausted as our construction goes on. In the second place, we will obtain $W_n \subset W_{n+1}$ for all $n \in \mathbb{N}$ ($\subset$ denoting proper inclusion here), provided that the elements of Δ have been arranged in a suitable way. This is easily proved by induction on the length of the indices of the various Γ's, observing that $\Phi_{\{\chi_0 \dots \chi_n\}} \subset \Phi_{\{\chi_0 \dots \chi_{n+1}\}}$ is definitely true.

It is evident now, that an enumeration $w: \mathbb{N}\backslash\{0\} \rightarrow \bigcup_{n \in \mathbb{N}} W_n$ exists, such that $w_1 = \Gamma_{\langle\rangle}$ and $W_n = \{w_1, \dots, w_{k_n}\}$ for k_n being the number of elements of W_n. To obtain a sequence of models, we define, for all $n \in \mathbb{N}$:

$M_n = \{1, \dots k_n\}$;

iR_nj iff for $w_i = \Gamma_p$, $w_j = \Gamma_q$, p is a proper initial segment of q;

$i \Vdash_{M_n} p$ iff $p \in w_i$.

Hereby obviously a sequence of models is defined. This sequence of models will from now on be referred to as the *L-canonical s.o.m. for* Δ (even though the sequence is not uniquely determined by this process). The clauses (a)-(f) are now easily proved. As to clause (g), we will prove: for all $\varphi \in \Phi_{\{\chi_0 \dots \chi_n\}}$, $n \in \mathbb{N}$:

$\forall i \in M_n \;\; \varphi \in w_i \Leftrightarrow i \Vdash_{M_n} \varphi$.

For atomic formulae this is clear from the definition of $\Vdash_{M_n}$. The cases $\varphi = \neg\psi$, $\varphi = \psi \wedge \chi$ are straightforward. Suppose $\varphi = \Box\psi$.

"$\Rightarrow$": if $\Box\varphi \in w_i$, then for all j such that iR_nj, $\varphi \in w_j$ holds, which is clear from the definition of R_n, hence by induction hypothesis: $j \Vdash_{M_n} \varphi$. This is exactly what is needed to conclude $i \Vdash_{M_n} \Box\varphi$.

"$\Leftarrow$": if $\Box\varphi \notin w_i$ and $\neg\Box\varphi \in \Phi_{\{\chi_0 \dots \chi_n\}}$, then $\neg\Box\varphi \in w_i$. There must be some j such that $w_j \in W_n$ and iR_nj and $\neg\varphi \in w_j$. But then $j \in M_n$

and, by induction hypothesis $j \Vdash_{M_n} \neg\varphi$, which is sufficient to conclude: $i \nVdash_{M_n} \Box\varphi$.

Now (g) is clear, for $\chi_m \in \Phi_{\{\chi_0 \dots \chi_n\}} \cap w_1$ for all m such that $0 \leqslant m \leqslant n$, so $1 \Vdash_{M_n} \chi_0 \wedge \dots \wedge \chi_n$. ⊠

3.5. Corollary. Let $\langle M_n \rangle_n$ be the L-canonical s.o.m. for Δ such that $\Delta \nvdash_L \bot$. A modal formula φ is called *stably true* in 1 if and only if a $m \in \mathbb{N}$ exists, such that for all $n \geqslant m$, $1 \Vdash_{M_n} \varphi$ holds. The following statement results from the construction of the L-canonical s.o.m. for Δ:

for all $\varphi \in \Phi_\Delta$, either φ or $\neg\varphi$ is stably true in 1.

Proof. Fix $\varphi \in \Phi_\Delta$. Apparently, $\varphi \in \Phi_{\{\chi_0 \dots \chi_m\}}$ for some $m \in \mathbb{N}$. Examining the proof of the foregoing lemma we can conclude: $\varphi \in w_1 \Leftrightarrow 1 \Vdash_{M_m} \varphi$ But the same will hold for all n such that $n \geqslant m$, since $\Phi_{\{\chi_0 \dots \chi_m\}} \subset \Phi_{\{\chi_0 \dots \chi_n\}}$. Thus, since either $\varphi \in w_1$ or $\neg\varphi \in w_1$, we can infer $\forall n \geqslant m\ 1 \Vdash_{M_n} \varphi$ or $\forall n \geqslant m\ 1 \Vdash_{M_n} \neg\varphi$. ⊠

Our next aim is to incorporate some modal syntax within Peano arithmetic.

3.6. Encoding modal logic. We will use some encoding of modal formulae as finite strings of symbols. Let in the following $\ulcorner X \urcorner$ be the numeral corresponding to the code of $X \in L_\Box$. We can extend the coding of symbols $\ulcorner\ \urcorner$ to the class of all well-formed formulae of $L_\Box$ by means of two primitive recursive functions, formalized in Peano arithmetic (extended with symbols for primitive recursive functions) as $imp(v_1,v_2)$ and $box(v_1)$, which satisfy the following:

$\ulcorner \varphi \rightarrow \psi \urcorner := imp(\ulcorner \varphi \urcorner, \ulcorner \psi \urcorner)$ and $\ulcorner \Box\varphi \urcorner := box(\ulcorner \varphi \urcorner)$.

We can use this encoding to formalize the interpretation function which assigns a sentence in the language of arithmetic to each modal formula as described in paragraph 1.8.

3.7. Formalizing interpretations. In order not to make things too illegible, we will restrict ourselves to the case where $L_\Box$ contains a single propositional letter p.

A binary function, formalized as $inter(v_1,v_2)$ exists, such that the following is provable in PA (using impl as the name of a two-place function in PA which gives the code of the implication of its arguments):

$$\forall v_1,v_2,x,y \ \mathrm{inter}(v_1,v_2) = \begin{cases} v_2 \text{ if } v_1 = \ulcorner p \urcorner; \\ \ulcorner 0=1 \urcorner \text{ if } v_1 = \ulcorner \bot \urcorner; \\ \mathrm{impl}(\mathrm{inter}(x,v_2),\mathrm{inter}(y,v_2)) \\ \quad \text{if } v_1 = \mathrm{imp}(x,y); \\ \ulcorner \exists p\, \mathrm{proof}(p,\mathrm{inter}(x,v_2)) \urcorner \text{ if } v_1 = \mathrm{box}(x); \\ 0 \text{ otherwise.} \end{cases}$$

By the representation theorem we are free to introduce this function, since it is clearly primitive recursive (assuming right bounds for imp and box). What inter(v_1,v_2) actually yields, is the code of the interpretation of a modal formula. The first variable ranges over codes of modal formulae and the second over codes of arithmetical expressions intended to replace the propositional letter p in an arithmetical interpretation.

In the following paragraphs we assume that certain properties of sequences of models as defined in 3.3 can be described by arithmetical expressions. Explicitly stated, this amounts to the following:

3.8. **Description of a sequence of models.** Let $\langle M_n \rangle_n$ be a sequence of models for a Δ such that $\Delta \nvdash_L \bot$. We will assume $\langle M_n \rangle_n$ to be described within PA arithmetic in the following sense:

(a) iR_nj is primitive recursive in i, j and n. It will be represented by formula $v_1R_{v_2}v_3$ of PA.

(b) k_n for $M_n=\{1,\dots k_n\}$ is primitive recursive in n. It will be represented by a function symbol k_{v_1} of PA.

(c) Assuming that $i \Vdash_{M_n} \ulcorner p \urcorner$ is primitive recursive, we can conclude that $i \Vdash_{M_n} \ulcorner \varphi \urcorner$ is primitive recursive in i, n and the code $\ulcorner \varphi \urcorner$ of φ (using (a) and (b)), since it can be defined by means of recursion in the code of φ and the truth-value of any of the subformulae of φ, which can be determined in a finite amount of steps. Thus, we assume it to be represented by a formula $v_1 \Vdash_{M_{v_2}} v_3$ of PA.

3.9.Definition. Let $\langle M_n \rangle_n$ be a s.o.m. for Δ such that $\Delta \nvdash_L \bot$. In case Δ is properly infinite, viz. $\Delta \vdash_L \neg\Box^n\bot$ for all $n \geqslant 1$, we can rearrange the elements of Δ in such a way, that $k_n > n$ for all $n \in \mathbb{N}$. We define a PRIM predicate rel' in the language of arithmetic as follows:

$$\mathrm{rel}'(v_1,v_2) := \neg\exists p,x,n \begin{bmatrix} p<v_1 \wedge x<v_1 \wedge n<v_1 \wedge \\ v_1 = k_n+1 \wedge 1 \Vdash_{M_n} x \wedge \\ \mathrm{proof}(p, \mathrm{impl}(\mathrm{inter}(x,v_2), \ulcorner 0=1 \urcorner)) \end{bmatrix}$$

The first variable in rel' is intended to range over all nodes of models in $\langle M_n \rangle_n$, the second over interpretations of the propositional letter p as in the definition of inter in paragraph 3.7. The purpose of this definition can be described as follows: let v_2 determine the interpretation of a set Δ of modal formulae in a language with one single propositional letter p. If Δ^* is inconsistent with PA, then some finite subset of Δ must be responsible for this inconsistency. Therefore some x being the modal code of a finite conjunction of elements of Δ must exist, such that $\mathrm{proof}(p, \ulcorner(\ulcorner * \mathrm{inter}(x,v_2) * \ulcorner \rightarrow \urcorner * \ulcorner 0=1 \urcorner * \ulcorner)\urcorner)$ holds for some p. But this finite conjunction has a model in $\langle M_n \rangle_n$. This really is a paradoxical situation. For each v_2, $\mathrm{rel}'(v_1,v_2)$ is true for those v_1 below which no such paradoxical situation can arise. These v_1 are the relevant nodes in the s.o.m.. We collect some facts about rel':

3.10.Lemma. Let Δ, $\langle M_n \rangle_n$ and rel' be as above and let F be a sentence of arithmetic and $(\ )^*$ an interpretation of modal formulae in the language $L_\Box = \{\bot, \rightarrow, (,), \Box\} \cup \{p\}$ which assigns F to p. Define: $\Delta^* = \{\chi^* \mid \chi \in \Delta\}$. The following holds:

If $PA + \Delta^* \vdash 0=1$, then for some $i \in \mathbb{N}$, $\mathbb{N} \models \neg\mathrm{rel}'(i, \ulcorner F \urcorner)$.

Proof. If $PA + \Delta^* \vdash 0=1$, then we do have a number p_0 coding the shortest proof of a sentence of the form $\chi_{j_1}^* \wedge \ldots \wedge \chi_{j_m}^* \rightarrow 0=1$ from the axioms of PA, so, $\mathrm{proof}(p_0, \langle \mathrm{inter}(x_0, \ulcorner F \urcorner), \ulcorner \rightarrow \urcorner, \ulcorner 0=1 \urcorner \rangle)$ holds for x_0 equal to $\ulcorner \chi_{j_1} \wedge \ldots \wedge \chi_{j_m} \urcorner$. Let n_0 by the definition of $\langle M_n \rangle_n$ be such that $\forall n \geqslant n_0\ 1 \Vdash_{M_n} \chi_{j_1} \wedge \ldots \wedge \chi_{j_m}$. This n_0 can be found primitive recursively, since $n_0 \leqslant \max\{j_1, \ldots, j_m\}$. As for all $n \in \mathbb{N}$ $k_n > n$, we can choose a $n \geqslant n_0$ such that $p_0 < k_n + 1 \wedge x_0 < k_n + 1$. This yields the desired result.

3.11. Embedding a sequence of models. Let $\langle M_n \rangle_n$ be a s.o.m. for a properly infinite Δ such that $\Delta \nvdash_L \bot$. We use the following abbreviations:

iRj for $\exists n(i \leqslant k_n \wedge j \leqslant k_n \wedge iR_n j)$;
$i\underline{R}j$ for $iRj \vee i=j$;
$i \circ j$ for $\neg(i\underline{R}j \vee j\underline{R}i)$;
$i \Vdash \ulcorner p \urcorner$ for $\exists n\ i \Vdash_{M_n} \ulcorner p \urcorner$.

Moreover, we assume R to be provably monotone, that is $PA \vdash \forall i,j(iRj \rightarrow j > i)$. This is just a matter of renumbering the nodes of the various M_n's in an orderly way. Let in the following N, neg and subst be formalizations of primitive recursive functions, such that the statements:

$N(k) = \ulcorner k \urcorner$;
$neg(\ulcorner \varphi \urcorner) = \ulcorner \neg \varphi \urcorner$;
$subst(\ulcorner A(v_1) \urcorner, \ulcorner t \urcorner) = \ulcorner A(t) \urcorner$

are provable in PA for k being any numeral, φ any sentence and $A(v_1)$ any predicate containing the free variable v_1 and t any term. For technical purposes we add an extra node 0 to the s.o.m. and define: $0 \Vdash p$ if and only if $1 \Vdash p$. Our equipment is now sufficiently developed to make the following definitions (using $Pr_{\neg}(p,v_3,j)$ short for $proof(p,neg(subst(v_3,N(j))))$).

$$J_1(v_1,v_2,v_3) := \forall j(v_1Rj \wedge rel'(j,v_2) \rightarrow \neg\exists p\, Pr_{\neg}(p,v_3,j)) \wedge \exists p\, Pr_{\neg}(p,v_3,v_1) \wedge rel'(v_1,v_2) \wedge \exists v_4 > v_1 \neg rel'(v_4,v_2);$$

$$J_2(v_1,v_2,v_3) := \forall j(v_1 \circ j \wedge rel'(j,v_2) \rightarrow \exists k[k \circ j \wedge rel'(k,v_2) \wedge k\underline{R}i \wedge \exists p\,(Pr_{\neg}(p,v_3,k) \wedge \neg\exists q \leqslant p\, Pr_{\neg}(p,v_3,j))]);$$

$$Lim_{>0}(v_1,v_2,v_3) := (v_1 = 1 \wedge J_1(v_1,v_2,v_3)) \vee (v_1 > 1 \wedge J_1(v_1,v_2,v_3) \wedge J_2(v_1,v_2,v_3));$$

$$Lim(v_1,v_2,v_3) := (v_1 = 0 \wedge \forall v_5(v_5 \neq 0 \rightarrow \neg Lim_{>0}(v_5,v_2,v_3))) \vee (v_1 \neq 0 \wedge Lim_{>0}(v_1,v_2,v_3)).$$

One instantly notices the similarity between these definitions and those of the sentences in 2.2. Apart from the variable v_2 (which only serves within the context of a fixed point definition as will be explained below), J_1 and J_2 resemble closely the schemes from which λ_1 and λ_i for $i > 1$ in 2.2 were drawn. The difference is, that the finite disjunctions and conjunctions are replaced by quantifiers ranging over the relevant nodes of the given s.o.m. only.

Let F be the formalization of a primitive recursive function, such that $F(\ulcorner\lambda(v_1)\urcorner)=\ulcorner\exists i[\lambda(i)\wedge i\Vdash \ulcorner p\urcorner]\urcorner$ is provable for every λ containing v_1 as a free variable. By the free variable-version of the fixed point theorem we obtain a formula $\lambda(v_1)$ of PA-arithmetic as a fixed point of the expression $Lim(v_1,F(v_3),v_3)$. Thus, we have:

$$PA\vdash\forall v_1(\lambda(v_1)\leftrightarrow Lim(v_1,\ulcorner\exists i[\lambda(i)\wedge i\Vdash \ulcorner p\urcorner]\urcorner,\ulcorner\lambda(v_1)\urcorner).$$

Unraveling the definitions, we can easily prove the following three clauses (writing $rel(i)$ for $rel'(i,F(\ulcorner\lambda(v_1)\urcorner))$ and $\Box\neg\lambda(i)$ short for $\exists p\,Pr_\neg(p,\ulcorner\lambda(v_1)\urcorner,i)$ and likewise in witness-comparison formulae):

(a) $PA\vdash\lambda(0)\leftrightarrow\forall i(i\neq0\to\neg\lambda(i))$;

(b) $PA\vdash\lambda(1)\leftrightarrow\Box\neg\lambda(1)\wedge\forall j(1Rj\wedge rel(j)\to\neg\Box\neg\lambda(j))$
$\wedge rel(1)\wedge\exists j>1\neg rel(j)$;

(c) $PA\vdash\forall i>1(\lambda(i)\leftrightarrow\Box\neg\lambda(i)\wedge\forall j(iRj\wedge rel(j)\to\neg\Box\neg\lambda(j))$
$\wedge\forall j(i\circ j\wedge rel(j)\to\exists k[k\circ j\wedge rel(k)\wedge k\underline{R}i\wedge$
$\Box\neg\lambda(k)\prec\Box\neg\lambda(j)])\wedge rel(i)\wedge\exists j>i\neg rel(j)$.

3.12. Consistency lemma. Let Δ be properly infinite and such that $\Delta\nvdash_L\bot$ and let $\langle M_n\rangle_n$ be a s.o.m. for Δ. An extra node 0 is added and $0\Vdash p$ is defined as equivalent to $1\Vdash p$. We invoke 3.11 to get a predicate $\lambda(v_1)$ satisfying clauses (a), (b) and (c) as above. We define an interpretation of all modal formulae by stipulating:

$$p^*:=\exists i(\lambda(i)\wedge i\Vdash \ulcorner p\urcorner).$$

This is a consistent interpretation of Δ.

Proof. Suppose that a conjunction $\chi_{j_1}\wedge\ldots\wedge\chi_{j_m}$ of elements of Δ would exist, such that $PA\vdash\neg(\chi_{j_1}^*\wedge\ldots\wedge\chi_{j_m}^*)$. By lemma 3.10 we can assume the existence of numbers i_0 and n_0 such that $n_0<i_0$ and $1\Vdash_{M_{n_0}}\chi_{j_1}\wedge\ldots\wedge\chi_{j_m}$, $k_{n_0}+1=i_0$ and $\neg rel'(i_0,\ulcorner p^*\urcorner)\wedge\forall i<i_0\, rel'(i,\ulcorner p^*\urcorner)$. Since the last expression is equivalent to a Δ_0-formula, we obtain: $PA\vdash\forall i(rel'(i,\ulcorner p^*\urcorner)\leftrightarrow i<i_0)$, hence $\forall i\geqslant i_0\neg\lambda(i)$ is a theorem of PA. We can therefore rewrite clauses (a), (b) and (c) as follows (writing R for R_{n_0}):

(a') $PA\vdash\lambda(0)\leftrightarrow\neg\bigvee_{1\leqslant i<i_0}\lambda(i)$

(b') $PA\vdash\lambda(1)\leftrightarrow\Box\neg\lambda(1)\wedge\bigwedge_{1<i<i_0}\neg\Box\neg\lambda(i)$

(c') $PA\vdash\lambda(i)\leftrightarrow\Box\neg\lambda(i)\wedge\bigwedge_{iRj}\neg\Box\neg\lambda(j)\wedge\bigwedge_{i\circ j}\bigvee_{\substack{k\underline{R}i\\ k\circ j}}\Box\neg\lambda(k)\prec\Box\neg\lambda(j)$

for all i such that $1<i<i_0$.

We will now concentrate on $M_{n_0} = \langle M_{n_0}, R_{n_0}, \Vdash_{M_{n_0}} \rangle$. Since we already had $M_{n_0}=\{1,\dots, k_{n_0}\}$, we know that, by the definition of rel', (a'), (b') and (c') involve exactly the nodes of M_{n_0}. This means that we can apply lemma 2.3 to the set of sentences $\{\lambda(i) | 0 \leqslant i \leqslant k_{n_0}\}$. To avoid confusion we define an interpretation ( )** as follows:

$$p^{**} := \bigvee_{i \Vdash_{M_{n_0}} p} \lambda(i)$$

It is evident, that $p^{**} \leftrightarrow p^*$ is a theorem of PA. By lemma 1.9 we can now conclude: $PA \vdash \neg(\chi_{j_1}^{**} \wedge \dots \wedge \chi_{j_m}^{**})$. But, on the other hand, since $1 \Vdash_{M_{n_0}} \chi_{j_1} \wedge \dots \wedge \chi_{j_m}$, we obtain $PA \vdash \lambda(1) \rightarrow \chi_{j_1}^{**} \wedge \dots \wedge \chi_{j_m}^{**}$ by the completeness theorem (cf. 2.7). Thus, $PA + \lambda(1)$ would be inconsistent, contradicting 3) of lemma 2.3. This completes our proof, because the assumption was apparently absurd. ⊠

As a matter of fact, something stronger than consistency can be obtained from the interpretation defined in 3.12. We will for that purpose invoke the following lemma, which may be considered as the relativised counterpart of lemma 2.3.

3.13. Lemma. Let $\lambda(v_1)$ for a given s.o.m. be defined as in 3.11. We will use the following abbreviation:

$c(n) := \neg rel(n) \wedge \forall n' < n\, rel(n)$.

Furthermore, we use the symbol n_0 as a formalization of a primitive recursive function which gives us the index number m of the model M_m such that $n = k_m + 1$ if c(n) is the case. The following statements are provable:

1) $PA \vdash \forall n[c(n) \rightarrow \exists! i < n\, \lambda(i)]$;
2) $\mathbb{N} \vDash \lambda(0)$;
3) $PA \vdash \forall n[c(n) \rightarrow (\lambda(0) \leftrightarrow \forall i(0 < i < n \rightarrow \neg\Box\neg\lambda(i)))]$;
4) $PA \vdash \forall n[c(n) \rightarrow \forall i(0 < i < n \wedge \lambda(i) \rightarrow \forall j(iR_{n_0}j \rightarrow \neg\Box\neg\lambda(j)))]$;
5) $PA \vdash \forall n[c(n) \rightarrow \forall i(0 < i < n \wedge \lambda(i) \rightarrow \Box\exists j\, iR_{n_0}j \wedge \lambda(j))]$.

Proof. If we compare these statements to those in lemma 2.3, we see that 1) and 2) correspond to 1) and 2) of lemma 2.3, so do 4) and 5). 3) occurs as a) in the proof of the same lemma. The proofs of 1), 3), 4) and 5) are essentially as before. The only difference consists in the fact that the numeral **n** determined by the cardinality of the Kripke-model which provides a bound to the whole process described in the proof of that lemma, is replaced by the term k_n here. All the iterated conjunctions and disjunctions in that proof are replaced by bounded quantifiers. One just has to

note, that it is provable in **PA** that every finite set of proofs contains a smallest one.
As to 2) we remark that this follows immediately from (b) and (c) at the end of 3.11. ⊠

Just as in the proof of Solovay's completeness theorem, we can apply this lemma to obtain:

3.14. Lemma. Let Δ be a properly infinite assumption set of modal formulae, consistent with respect to **L**. Let $\langle M_n\rangle_n$ be a s.o.m. for this Δ, with an additional node 0 for which the forcing relation is extended in the usual way and let $\lambda(v_1)$ be defined as in 3.11. We define an interpretation of all modal formulae as in the consistency lemma, by stipulating:

$p^* := \exists i[\lambda(i) \wedge i \Vdash \ulcorner p \urcorner]$.

We already know, by 3.12, that $(\ )^*$ is a consistent interpretation of Δ. The following statement holds:

for all modal formulae φ:

$\mathrm{PA} \vdash \forall n[c(n) \rightarrow \forall i(1 \leqslant i < n \wedge \lambda(i) \rightarrow (\varphi^* \leftrightarrow i \Vdash_{M_{n_0}} \ulcorner \varphi \urcorner))]$.

Proof. By induction on the length of φ:
$\varphi = p$. This case is clear by the definition of p^* and 1) from lemma 3.13. So is the case where $\varphi = \bot$.
The case where $\varphi = \psi \rightarrow \chi$ is straightforward by induction. The difficult case is $\varphi = \Box\psi$.
"$i \Vdash_{M_{n_0}} \ulcorner \Box\psi \urcorner$": By the definition of $\Vdash$ we have:

$$\mathrm{PA} \vdash \forall n[c(n) \rightarrow \forall i(1 \leqslant i < n \wedge i \Vdash_{M_{n_0}} \ulcorner \Box\psi \urcorner \rightarrow \forall j(iR_{n_0}j \rightarrow j \Vdash_{M_{n_0}} \ulcorner \psi \urcorner))].$$

Thus, by induction hypothesis:

$\mathrm{PA} \vdash \forall n[c(n) \rightarrow \forall i(1 \leqslant i < n \wedge i \Vdash_{M_{n_0}} \ulcorner \Box\psi \urcorner \wedge \exists j\, iR_{n_0}j \wedge \lambda(j) \rightarrow \psi^*)]$,

hence

$\mathrm{PA} \vdash \forall n[c(n) \rightarrow \forall i(1 \leqslant i < n \wedge i \Vdash_{M_{n_0}} \ulcorner \Box\psi \urcorner \wedge \Box[\exists j\, iR_{n_0}j \wedge \lambda(j)] \rightarrow \Box\psi^*)]$,

applying 5), we now obtain:

$\mathrm{PA} \vdash \forall n[c(n) \rightarrow \forall i(1 \leqslant i < n \wedge i \Vdash_{M_{n_0}} \ulcorner \Box\psi \urcorner \wedge \lambda(i) \rightarrow \Box\psi^*)]$.

"$\neg i \Vdash_{M_{n_0}} \ulcorner \Box\psi \urcorner$": By the definition of $\Vdash$ we have:

$$\mathrm{PA} \vdash \forall n[c(n) \rightarrow \forall i(1 \leqslant i < n \wedge \neg i \Vdash_{M_{n_0}} \ulcorner \Box\psi \urcorner \rightarrow \exists j(iR_{n_0}j \wedge \neg j \Vdash_{M_{n_0}} \ulcorner \psi \urcorner))].$$

By induction hypothesis, we have:

$\mathrm{PA} \vdash \forall n[c(n) \rightarrow \forall j(1 \leqslant j < n \wedge \psi^* \wedge \neg j \Vdash_{M_{n_0}} \ulcorner \psi \urcorner \rightarrow \neg\lambda(j))]$,

hence, as the other formulae involved are Δ_0,

$\mathrm{PA} \vdash \forall n[c(n) \rightarrow \forall j(1 \leqslant j < n \wedge \Box\psi^* \wedge \neg j \Vdash_{M_{n_0}} \ulcorner \psi \urcorner \rightarrow \Box\neg\lambda(j))]$.

Combined with our first remark and 4) of lemma 3.13, this yields:

$$PA \vdash \forall n[c(n) \rightarrow \forall i(1 \leqslant i < n \wedge \neg i \Vdash_{M_{n0}} \ulcorner \Box\psi \urcorner \wedge \Box\psi^* \rightarrow \neg\lambda(i))].$$

Our proof of lemma 3.14 is now completed. ⊠

We will now direct our attention towards Solovay's second completeness theorem (cf. 1.11). The most interesting implication of this theorem is obviously the one which states that for every φ such that $S+\varphi$ is consistent, that is $\neg\varphi \notin S$, there is an interpretation ( )* which makes φ a true sentence of arithmetic. In order to extend the notion of S-consistency to infinite assumption sets, we give the following definition:

3.15. Definition. Let, for any assumption set Δ of modal formulae, Δ_S be defined as the smallest extension of Δ containing all formulae of the form $\Box\varphi \rightarrow \varphi$ for $\Box\varphi \in \Phi_\Delta$ and closed under modus ponens. Δ is called S-consistent if and only if $\Delta_S \nvdash_L \perp$.

It will be argued below that, under certain conditions, the interpretation we used in the consistency lemma is not only a consistent one, but even permits us to take the standard model of arithmetic as a model for the entire interpreted set. So, for a certain class of assumption sets we can strengthen the second completeness theorem to: if Δ is S-consistent, then an interpretation exists, such that $\mathbb{N} \vDash \varphi^*$ for all φ in Δ. Unfortunately, it is at this point unclear for which type of infinite assumption set these conditions can be fulfilled. In particular, this applies to condition (c) of the next theorem.

3.16. A larger case of truth. Let in the following Δ be S-consistent and properly infinite. we will assume that a s.o.m. $\langle M_n \rangle_n$ for Δ exists, with the following properties:

(a) $\langle M_n \rangle_n$ is stable for all subformulae of formulae in Δ, so: for all $\varphi \in \Phi_\Delta$, either φ or $\neg\varphi$ is stably true in 1;

(b) for all $\Box\varphi \in \Phi_\Delta$, $\Box\varphi \rightarrow \varphi$ is stably true in 1;

(c) relevant stability is provable, that is: for all $\varphi \in \Phi_\Delta$, if $\forall n' \geqslant n\ 1 \Vdash_{M_{n'}} \varphi$, then $PA \vdash \forall n' \geqslant n\ 1 \Vdash_{M_{n'}} \ulcorner \varphi \urcorner$.

Now let 0 be an additional node to $\langle M_n \rangle_n$ and the forcing relation for 0 defined as usual and let ( )* be defined by:

$p^* := \exists i[\lambda(i) \wedge i \Vdash p]$. The following holds:

for all $\varphi \in \Phi_\Delta$:

$\exists n \forall n' \geqslant n\ 1 \Vdash_{M_{n'}} \varphi \Rightarrow PA \vdash \lambda(0) \rightarrow \varphi^*$ and

$\exists n \forall n' \geqslant n\ 1 \Vdash_{M_{n'}} \neg\varphi \Rightarrow PA \vdash \lambda(0) \rightarrow \neg\varphi^*$.

Proof. The cases where $\varphi = \bot$ or $\varphi = \psi \rightarrow \chi$ or $\varphi = p$ are easily proved, using induction and stability.
Suppose $\varphi = \Box\psi \in \Phi_\Delta$ and $\exists n \forall n' \geqslant n\ 1 \Vdash_{M_{n'}} \Box\psi$. By our assumptions we have: $\exists n \forall n' \geqslant n\ 1 \Vdash_{M_{n'}} \psi$, so, by induction hypothesis: $PA \vdash \lambda(0) \rightarrow \psi^*$. But, since also $PA \vdash \forall n' \geqslant n \forall i\ 1 R i\ i \Vdash_{M_{n'}} \ulcorner \psi \urcorner$ follows from our assumptions, we obtain:
$PA \vdash \forall n[c(n) \rightarrow \forall i(1 \leqslant i < n \wedge \lambda(i) \rightarrow \psi^*)]$, so $PA \vdash \exists n(c(n)) \rightarrow \psi^*$. But since $PA \vdash \neg\lambda(0) \rightarrow \exists n(c(n))$ follows from the definition of $\lambda(0)$, we obtain $PA \vdash \psi^*$, so evidently $PA \vdash \lambda(0) \rightarrow \Box\psi^*$.
Now suppose that $\Box\psi$ is stably false in 1. By the definition of rel we can conclude: $PA \vdash \Box\psi^* \rightarrow \exists n(c(n))$. By lemma 3.14 we obtain: $PA \vdash \Box\psi^* \rightarrow \neg\lambda(1)$. Applying formalized Σ-completeness, this yields: $PA \vdash \Box\psi^* \rightarrow \Box\neg\lambda(1)$, so combining this with lemma 3.13, clause 3), we obtain $PA \vdash \Box\psi^* \rightarrow \neg\lambda(0)$. This completes our proof.
As a direct consequence of this proof, we obtain (since $\mathbb{N} \models \lambda(0)$):

$\mathbb{N} \models \varphi^*$ for all $\varphi \in \Delta$. ⊠

As we have already pointed out, this result cannot be extended straightaway to the class of all assumption sets Δ which are S-consistent. Although for each S-consistent Δ a s.o.m. $\langle M_n \rangle_n$ exists which has properties (a) and (b), we can take the L-canonical s.o.m. for Δ_S for this purpose, this is not immediately clear as to property (c). The reason why relevant stability might not be provable in a canonical s.o.m., is that the construction of this type of sequence involves the use of maximal consistent sets of formulae. On the other hand, we can, for certain assumption sets, use a s.o.m. which is considerably smaller than the canonical one. A fine example is provided by the assumption set defined in 1.7.

3.17. Example. Let a sequence $\langle \varphi_n \rangle_n$ of modal formulae be defined as in 1.7. Define an assumption set $\Delta = \{\chi_0, \chi_1, \chi_2, \ldots\}$ as follows:

$\chi_0 = \Diamond\Diamond\varphi_0$;

$\chi_{i+1} = \Box(\Diamond\varphi_i \rightarrow \Diamond\Diamond\varphi_{i+1})$.

As we have already seen, Δ is consistent with respect to L. It will be clear that Δ is even S-consistent. Now let a sequence

$\langle M_n \rangle_n$ be defined as follows (using $f(\langle i,j \rangle)=i(i+1)/2+j+1$ as a standard enumeration of ordered pairs):

$M_n := \{f(\langle x,y \rangle) | x,y \in \mathbb{N} \wedge y \leqslant x \wedge x \leqslant n+2\}$;

$f(\langle x,y \rangle) R_n f(\langle x',y' \rangle) \Leftrightarrow (y=0 \wedge x < x') \vee (x=x' \wedge y < y')$;

$f(\langle x,y \rangle) \Vdash_{M_n} p \Leftrightarrow y=1$.

It is easily verified that this defines a s.o.m. for Δ. Arguing as in the consistency lemma, we obtain a formula $\lambda(v_1)$, inducing a consistent interpretation $(\ )^*$ for this Δ. Since all relevant properties of this s.o.m. can be described by quite simple predicates, we may safely assume that relevant stability is provable. This is the case since the formulae χ_i have a very simple uniform shape and for all subformulae of χ_i it is exactly clear at which nodes they are forced and at which nodes they are not. So we can conclude: $\mathbb{N} \models \varphi^*$ for all $\varphi \in \Delta$ (using 3.16). Another interesting feature of this s.o.m. is, that we can prove that $PA+\lambda(1)$ is consistent. We reason as follows:

$PA \vdash \forall n[c(n) \to \forall i(1 \leqslant i < n \to (i=1 \leftrightarrow i \Vdash_{M_{n_0}} \ulcorner \Diamond\Diamond\varphi_0 \urcorner))]$. This is immediate from the construction of the s.o.m., so using lemma 3.14, we obtain: $PA \vdash \forall n[c(n) \wedge \neg\lambda(1) \to \lambda(0) \vee \Box\Box\neg\varphi_0^*]$. Suppose that $\neg\lambda(1)$ were a theorem of PA, then $\Box\neg\lambda(1)$ would be a theorem of PA as well, so, using lemma 3.13, we would obtain: $PA \vdash \forall n(c(n) \to \Box\Box\neg\varphi_0^*)$. But, since $\Diamond\Diamond\varphi_0 \in \Delta$, this yields: $PA \vdash \Box\Box\Box\neg\varphi_0^* \to \Box\Box\neg\varphi_0^*$, whence follows, by Löb's rule: $PA \vdash \Box\Box\neg\varphi_0^*$, contradicting the consistency of Δ^*. ⊠

Bibliography:

Artyomov, S.N., 'Arithmetically complete modal theories', Semiotics and information science 14 (Russian), Akad. Nauk, SSSR, Vsesojuz. Inst. Naucn. i. Tehn. Informaticii, Moscow, 1980.
Guaspari, D. and R.M. Solovay, 'Rosser Sentences', Annals of mathematical logic 16, pp. 81-99, 1979.
de Jongh, D.H.J., 'A Simplification of a Completeness Proof of Guaspari and Solovay', Studia Logica 46, pp. 187-192, 1987.
Smoryński, C., Self-reference and modal logic, Springer, New York, 1985.
Solovay, R.M., 'Provability interpretations of modal logic', Israel Journal of Mathematics, vol.25, pp. 287-304, 1976.
Visser, A., Aspects of Diagonalization and Provability, dissertation, Utrecht, 1981.
Visser, A., 'A descending hierarchy of reflection principles', Logic group preprint series No. 34, Utrecht, 1988.

The ITLI Prepublication Series