Extending ILM with an operator for Σ_1-ness[1]

by

Evan Goris

Abstract

In this paper we formulate a logic ΣILM. This logic extends ILM and contains a new unary modal operator Σ_1. The formulas of this logic can be evaluated on Veltman frames. We show that ΣILM is modally sound and complete with respect to a certain class of Veltman frames. An arithmetical interpretation of the modal formulas can be obtained by reading the Σ_1 operator as formalized Σ_1-ness in PA and $\rhd$ as formalized Π_1-conservativity between finite extensions of PA. We show that under this arithmetically interpretation ΣILM is sound and complete.

The main motivation for formulating ΣILM at all is that one counterexample for interpolation in ILM seems to emerge because of the lack of ILM to express Σ_1-ness. We show that ΣILM does not have interpolation either. Our counterexample seems to emerge because of the inability of ΣILM to express Σ-interpolation[7]. (A formula $\phi \to \psi$ has a Σ_1-interpolant if there exist some $\sigma \in \Sigma_1$ such that PA $\vdash \phi \to \sigma$ and PA $\vdash \sigma \to \psi$.)

[1]The text of this paper formed the master's thesis of the author at the ILLC, June 2003, under supervision of Prof. Dr. D.H.J. de Jongh.

Contents

In this paper the logic ΣILM is introduced and some of its properties investigated. ΣILM is basically the union of the known logic ILM and the in this paper introduced logic ΣL. In section 1 the main preliminaries are explained. In section 2 and 3 the logics ΣL and ΣILM are treated. We show that both logics are sound and complete w.r.t. a modal and an arithmetical interpretation. We also show that both logics lack the interpolation property.

1 Introduction

In section 1.1 the modal logic ILM and its relation to arithmetic is introduced. We define a class of structures called Veltman frames. These serve as a basis for a modal semantics for all the logics we will see in this paper. More on ILM will be introduced in later sections when needed. In section 1.2 we motivate our study.

1.1 Preliminaries: Interpretability logics

In this paper we will be concerned with what are known as interpretability logics. These are nonstandard extensions of the normal modal logic GL and their language contains, besides the $\Box$, a binary modal operator $\rhd$.

Definition 1.1 (PROP). *PROP* is a (fixed) countable infinite set of propositional variables.

Definition 1.2 (IL-formulas). *IL-formulas* are built up using PROP, the propositional connectives, a unary modal operator $\Box$ and a binary modal operator $\rhd$.

With regard to priorities $\rhd$ behaves similarly as $\to$, although $\rhd$ binds stronger than $\to$. So $A \wedge B \rhd C$ means $(A \wedge B) \rhd C$ and $A \to B \rhd C$ means $A \to (B \rhd C)$.

As is well-known one can give arithmetical meaning to modal formulas by substituting for the proposition variables (arbitrary) arithmetical sentences and interpreting the $\Box$ as a formalization of 'provable in PA'. See for example [2]. We can extend this to formulas which contain $\rhd$ as follows. If A and B are modal formulas and A^* and B^* arithmetical ones, the 'arithmetical meaning' of A and B respectively, then the arithmetical meaning of $A \rhd B$ is a formalization of:

$$\text{PA} + A^* \text{ interprets } \text{PA} + B^*.$$

In general: a theory T interprets a theory S if there exists a translation of formulas of S into formulas of T such that T proves all the (translations of the) theorems of S (for a precise formulation see [10][19], I will not bother with this here since we will switch to another interpretation of $\rhd$ anyway).

Why is interpreting (S into T) interesting? Interpreting a theory into another is useful for e.g. showing relative consistency or, as we shall see, showing partial conservation results. Moreover, if we replace PA by some other theory T (that is: $\Box$ means provability in T and $\rhd$ means interpretability between finite extensions of T) then $\rhd$ gives us a means to distinguish between different theories T far better than was possible with only the $\Box$. For example $I\Sigma_1$ and PA are indistinguishable using $\Box$ only but $A \rhd B \to \Box(A \rhd B)$ is a valid principle (what this means will be made precise below) in $I\Sigma_1$ but not in PA. So, in general, replacing PA by another theory will change the discussion. However there is a basic system with nice properties that is present in all (reasonable) choices. This logic is called IL. It should be noted that this is not the largest logic present in all reasonable choices. What is, is still open. For conjectures (and for a definition of reasonable) see [13].

Definition 1.3 (IL). With IL we will refer to the following set of axiom schemata:

1. $\Box(A \to B) \to \Box A \to \Box B$,
2. $\Box(\Box A \to A) \to \Box A$,
3. $\Box(A \to B) \to (A \rhd B)$,
4. $(A \rhd B) \wedge (B \rhd C) \to (A \rhd C)$,
5. $(A \rhd C) \wedge (B \rhd C) \to (A \vee B \rhd C)$,
6. $(A \rhd B) \to (\Diamond A \to \Diamond B)$,
7. $\Diamond A \rhd A$.

We will obtain *the logic IL* by taking all instances of the above schemata, classical propositional logic in the enriched language[2], and close off under necessitation and modus ponens. We write $IL \vdash A$ for $A \in$ *the logic IL*. Without danger of confusion we speak of IL when we mean *the logic IL*.

The class of valid interpretability principles for certain theories can be axiomatized by adjoining to IL appropriate axiom schemata. PA is a theory for which such a schema has been obtained. The schema called (M).

Definition 1.4 (ILM ,(M)). With (M) we denote the schema:

$$A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C.$$

ILM is the set of schemata $\{$*(M)*$\}$ + IL and we obtain *the logic ILM* by taking the universal closure of (the schemata) ILM and close off under necessitation and modus ponens. Again we write ILM when we mean *the logic ILM*.

We can evaluate IL-formulas on Veltman frames and in Veltman models.

Definition 1.5 (Veltman Frame). A *Veltman frame*, or just *frame*, is a triple $F = \langle W, R, S\rangle$ where

1. W is a set, the domain of F,

[2] Alternatively one can add a few schemata which axiomatize classical propositional logic.

2. R is a binary relation on W,

3. S is a ternary relation on W such that for all w, a, b: $(w, a, b) \in S \Rightarrow (w, a), (w, b) \in R$.

We will write aS_wb for $(w, a, b) \in S$, and S_w designates the binary relation $\{(a, b) \mid (w, a, b) \in S\}$.

Definition 1.6 (Veltman model). A *Veltman model*, or simply *model*, is a quadruple $M = \langle W, R, S, V\rangle$ where $\langle W, R, S\rangle$ is a Veltman frame and V is a function PROP $\longrightarrow \mathcal{P}(W)$. With a valuation appropriate for a frame F we mean such a function PROP $\longrightarrow \mathcal{P}(W)$.

Veltman frames and Veltman models will serve as a basis for a semantics for all logics to be seen in this paper. How exactly will be postponed until later sections.

1.2 Motivation

Although modally and arithmetically complete (for PA) there still is a problem with ILM. It does not have the interpolation property. This problem is easier addressed when we switch to another arithmetical interpretation of $\rhd$.

Definition 1.7 (Bounded quantifier, $\Pi_1!$, $\Sigma_1!$, Π_1 and Σ_1-formulas). $\forall x{\leq}z\, \phi(x)$ abbreviates $\forall x(x{\leq}z \rightarrow \phi(x))$ and $\exists x{\leq}z\, \phi(x)$ abbreviates $\exists x(x{\leq}z \wedge \phi(x))$. $\forall x{\leq}z$ and $\exists x{\leq}z$ are called *bounded quantifiers*.

A $\Pi_1!$-*formula* is a formula of the form $\forall x\phi(x)$. Where in $\phi(x)$ all quantifiers occur bounded. A $\Sigma_1!$-*formula* is a formula of the form $\exists x\phi(x)$. Where in $\phi(x)$ all quantifiers occur bounded. A Π_1-*formula* is a formula equivalent to a $\Pi_1!$ *formula*. A Σ_1-*formula* is the negation of a Π_1 *formula*.

Definition 1.8 (Π_1-Conservativity). Let T and S be theories. We say that S is Π_1-*Conservative* over T if for any $\Pi_1!$ sentence π

$$S \vdash \pi \Rightarrow T \vdash \pi.$$

Besides being the logic of interpretability, ILM happens to be the logic of Π_1-Conservativity of PA as well (in fact in PA these notions coincide, see e.g. [10])[3].

Now let us see why the (M) schema is, in some sense, true when we read $\rhd$ as Π_1-Conservativity between PA finite extensions of PA and $\Box$ as provability in PA. We show that for all arithmetical first-order formulas ϕ, ψ, η:

$$\phi \rhd \psi \Rightarrow \phi \wedge \Box\eta \rhd \psi \wedge \Box\eta. \tag{1}$$

Suppose $\phi \rhd \psi$ and π is some Π_1 sentence provable in PA $+ \psi \wedge \Box\eta$. Then PA $+ \psi$ proves the Π_1 sentence (Π_1 since $\Box\eta$ is a Σ_1 statement) $\Box\eta \rightarrow \pi$, and thus PA $+ \phi$ proves $\Box\eta \rightarrow \pi$ as well, conclusion: PA $+ \phi \wedge \Box\eta$ proves π.

[3] It should be noted however that ILM is the logic of Π_1-Conservativity for $I\Sigma_1$ as well. And thus we can no longer distinguish between the two under this interpretation of $\rhd$.

Now let us consider a well-known counterexample, due to Ignatiev, for interpolation in ILM [19][4]:

$$\Box(p \leftrightarrow \Box q) \to (r \rhd s \to r \wedge p \rhd s \wedge p). \tag{2}$$

Suppose we could express Σ_1-ness by a ILM-formula with one proposition variable, say $\Sigma_1(p)$. The above proof that the (M) schema is true actually shows that the following schema is true:

$$\Sigma_1(C) \to (A \rhd B \to A \wedge C \rhd B \wedge C).$$

Moreover that argument can be carried out in PA and that means, by definition, that (1) is arithmetically valid. Since $\Box(A \leftrightarrow \Box C) \to \Sigma_1(A)$ is arithmetically valid as well, and ILM is known to prove all arithmetically valid formulas, we would have an interpolant for (2) (namely $\Sigma_1(p)$).

The main subject of this paper is to investigate the possibility of adjoining an operator to ILM which should express Σ_1-ness and see what it gives us. As a starting point we reduce the logic HGL [8], which contains, among other things, for each $n \geq 1$ a predicate expressing Σ_n-ness, to a logic, which we will call ΣL, with only the Σ_1 predicate. We give a simple proof of modal and arithmetical completeness for this logic. It turns out, however, that ΣL has no interpolation. These results are extended to show that it is insufficient to extend ILM with a Σ_1 predicate in order to obtain a logic with interpolation.

1.3 Notations

In this section we agree on some notations and conventions.

Upper case characters $A, B, C, \ldots$ range over modal formulas (of all kinds to be seen in this paper). The lower case characters $a, b, c, \ldots, p, q, r, \ldots$ range over elements of PROP and over nodes in frames and models.

For models M we will use the notation M for both the model and its domain. Similarly for frames. If $F = \langle W, R, S\rangle$ then we write W^F for W, R^F for R, S^F for S:

$$\langle W^F, R^F, S^F\rangle =_{\text{def}} F.$$

We define all the set-theoretic operations and tests on frames by performing them on their components. So for instance:

$$F_0 \cap F_1 =_{\text{def}} \langle W^{F_0} \cap W^{F_1}, R^{F_0} \cap R^{F_1}, S^{F_0} \cap S^{F_1}\rangle.$$

Similar definitions hold for $F_0 \cup F_1$ and $F_0 \subseteq F_1$. For two models M_0 and M_1 similar conventions hold but of course this only makes sense when $x \in M_0 \cap M_1 \Rightarrow \forall p \in \text{PROP} : x \in V^{M_0}(p) \Leftrightarrow x \in V^{M_1}(p)$.
For binary relations R we write R^* for the reflexive transitive closure of R.

[4]In [19] it is shown that $\Box(p \leftrightarrow \Box q) \to (r \rhd s \to \Diamond r \wedge p \rhd s \wedge p)$ is a counterexample for interpolation. The proof works unmodified for $\Box(p \leftrightarrow \Box q) \to (r \rhd s \to r \wedge p \rhd s \wedge p)$,the original unpublished counterexample by Ignatiev, as well.

If A is a modal formula then

$$\boxdot A =_{\text{def}} \Box A \wedge A.$$

If Γ is a finite set of formulas (first order or modal), say $\Gamma = \{\gamma_0, \ldots, \gamma_{n-1}\}$ then

$$\bigwedge \Gamma =_{\text{def}} \gamma_0 \wedge \cdots \wedge \gamma_{n-1}$$

and

$$\bigvee \Gamma =_{\text{def}} \gamma_0 \vee \cdots \vee \gamma_{n-1}.$$

If Γ is a set of modal formulas then

$$\Box\Gamma =_{\text{def}} \{\Box\gamma \mid \gamma \in \Gamma\}$$

and

$$\boxdot\Gamma =_{\text{def}} \{\boxdot\gamma \mid \gamma \in \Gamma\}.$$

2 The logic ΣL

A a modal logic called ΣL is defined. We show how the formulas of this logic can be evaluated on Veltman frames. A specific class of Veltman frames is defined. These are essentially the frames Japaridze used for his HGL [8] reduced to the Σ_1-case. In Section 2.2 and 2.3 ΣL is shown, by a relatively simple proof, to be sound and complete w.r.t. this class of frames. In Section 2.4 we show that ΣL does not have interpolation. In Section 2.5 we make a note on an arithmetical interpretation of ΣL but a treatment is postponed until Section 4.2.

2.1 Definitions

Definition 2.1 (ΣL-formulas). With a *ΣL-formula* we will mean a formula over PROP constructed using boolean connectives and two unary modal operators: $\Box$ and Σ_1.

Σ_1 binds like the $\Box$. So e.g. $\Sigma_1 A \to B$ means $(\Sigma_1 A) \to B$. We will be evaluating ΣL-formulas on Veltman frames.

Definition 2.2 (ΣL forcing relation). For a model $M = \langle W, R, S, V\rangle$ we let $\models_M$ be *the unique relation between elements w from W and* ΣL*-formulas* satisfying

1. $w \models_M p$ if $w \in V(p)$,
2. $w \models_M \Box A$ if for each v s.t. wRv: $v \models_M A$,
3. $w \models_M \Sigma_1 A$ if for each u, v s.t. $uS_w v$: $u \models_M A \Rightarrow v \models_M A$,
4. the usual constraints for boolean connectives.

In what follows we write $M, v \models A$ for $v \models_M A$ or, when M is clear from context, $v \models A$.

For a justification for this forcing relation one can think about the following. In our setting Σ_1 sentences are those sentences that are preserved along the S_w relations. By the Lös-Tarski theorem (see [6]) and Matjasevich's theorem (see [16]) a sentence is Σ_1 in PA precisely if it is preserved along embeddings of models of PA. For a full discussion on this see the appendix in [19].

Definition 2.3 (Frame validity). We say that a formula A is *valid on a frame* $F = \langle W, R, S\rangle$, and write $F \models A$, whenever for any valuation $V : \mathrm{PROP} \longrightarrow \mathcal{P}(W)$ and any $w \in W$, $w \models_{\langle W,R,S,V\rangle} A$.

Definition 2.4 (ΣL). With ΣL we denote the set of schemata:

1. $\Box(A \to B) \to (\Box A \to \Box B)$,
2. $\Box(\Box A \to A) \to \Box A$,
3. $\Sigma_1 A \wedge \Sigma_1 B \to \Sigma_1(A \wedge B)$,
4. $\Sigma_1 A \wedge \Sigma_1 B \to \Sigma_1(A \vee B)$,
5. $\Sigma_1 A \wedge \Box(A \leftrightarrow B) \to \Sigma_1 B$,
6. $\Sigma_1 A \to \Box\Sigma_1 A$,
7. $\Sigma_1 \bot$,
8. $\Sigma_1 \Box A$,
9. $\Sigma_1 \Sigma_1 A$,
10. $\Sigma_1 A \to \Box(A \to \Box A)$.

The logic ΣL is obtained by taking the universal closure of the above schemata, classical propositional logic and closing off under necessitation and modus ponens. We write $\Sigma L \vdash A$, or $\vdash A$, for $A \in$ *the logic* ΣL.

Definition 2.5 (ΣL-frame). $F = \langle W, R, S\rangle$ is a ΣL*-frame* if F is a frame and

1. R is transitive and conversely well-founded,
2. for all a, b, c, w, t:
   (a) $aS_w bRc \Rightarrow aRc$,
   (b) $wRaRb \Rightarrow aS_w b$,
   (c) wRv and $aS_v b \Rightarrow aS_w b$,
   (d) $wS_t v$ and $aS_v b \Rightarrow aS_w b$.

It is possible to take the S_w's to be transitive and reflexive. In later sections we will do so but for now we keep it like this.

Definition 2.6 (ΣL-model). A model $M = \langle W, R, S, V\rangle$ is a ΣL*-model* if $\langle W, R, S\rangle$ is a ΣL-frame

2.2 Modal soundness

In this section we will show that the logic $\Sigma\mathrm{L}$ is sound with respect to $\Sigma\mathrm{L}$-frames.

Theorem 2.7. *If* $\Sigma\mathrm{L} \vdash A$ *then* A *is valid on each* $\Sigma\mathrm{L}$*-frame.*

Proof. It is well-known that frame validity is preserved under modus ponens (if A and $A \to B$ are valid on F then so is B) and necessitation (if A is valid on F then so is $\Box A$). See for example [5]. Propositional tautologies are clearly valid on any frame. What is left is to show that any formula which is of one of the forms 1-10 from Definition 2.4 is valid on any $\Sigma\mathrm{L}$-frame.

The (relevant parts of the) proof of modal soundness for GL (see [2]) with respect to transitive conversely well-founded Kripke frames can be copied here to show Items 1 and 2.

So we show the Cases 3-10. In each of the cases below let F be a $\Sigma\mathrm{L}$-frame, $w \in F$ and let V be a valuation for F (a function PROP $\longrightarrow \mathcal{P}(W^F)$).

3. Suppose $w \models \Sigma_1 A \wedge \Sigma_1 B$. Let x, y be such that wRx, y and $xS_w y$ and $x \models A \wedge B$. Then $x \models A$, so $y \models A$, and $x \models B$, so $y \models B$, and therefore $y \models A \wedge B$.

4. Suppose $w \models \Sigma_1 A \wedge \Sigma_1 B$. Let x, y be such that wRx, y and $xS_w y$ and $x \models A \vee B$. If $x \models A$ then $y \models A$, and if $x \models B$ then $y \models B$. Either way $y \models A \vee B$.

5. Suppose $w \models \Sigma_1 A \wedge \Box(A \leftrightarrow B)$. Let x, y be such that wRx, y, $xS_w y$ and $x \models B$. Then $x \models A$, so $y \models A$, and thus $y \models B$.

6. Suppose $w \models \Sigma_1 A$. Suppose wRv, we will show $v \models \Sigma_1 A$. Choose x, y such that vRx, y and $xS_v y$ and $x \models A$. By Property 2c of $\Sigma\mathrm{L}$-frames we have $xS_w y$ and therefore $y \models A$.

7. Clear, since the situation $x \models \perp$ cannot occur.

8. Choose x, y such that $wRxS_w y$, $x \models \Box A$. We want to show: $y \models \Box A$. Choose z such that yRz. By Property 2a of $\Sigma\mathrm{L}$-frames (Definition 2.5) we have xRz and thus $z \models A$.

9. Choose x, y such that wRx, y and $xS_w y$ and $x \models \Sigma_1 A$. We want to show $y \models \Sigma_1 A$. Choose z_0, z_1 such that yRz_0, z_1 and $z_0 S_y z_1$ and $z_0 \models A$. By Property 2d of $\Sigma\mathrm{L}$-frames we have $z_0 S_x z_1$ and thus $z_1 \models A$.

10. Suppose $w \models \Sigma_1 A$. Choose v such that wRv. We want to show $v \models A \to \Box A$. So suppose that $v \models A$. If u is such that vRu then by Property 2b of $\Sigma\mathrm{L}$-frames we have $vS_w u$ and thus $u \models A$. $\dashv$

2.3 Modal completeness

2.3.1 Introduction and definitions

In this section we prove completeness of the logic $\Sigma\mathrm{L}$ with respect to $\Sigma\mathrm{L}$-frames.

Theorem 2.8 (Completeness). *Let* A *be a* $\Sigma\mathrm{L}$*-formula such that* $\Sigma\mathrm{L} \nvdash A$. *Then there exist a* $\Sigma\mathrm{L}$*-model* $M = \langle W, R, S, V\rangle$ *and* $w \in W$ *such that* $M, w \not\models A$.

The proof will be postponed until Section 2.3.3. The main ingredient in the proof is the notion of a maximal consistent set, MCS for short.

Definition 2.9 (Inconsistent/Consistent). A set of formulas Γ is called *inconsistent* if for some finite subset $\Gamma' \subseteq \Gamma$ we have $\Sigma\mathrm{L} \vdash \bigwedge \Gamma' \to \bot$. A set of formulas is *consistent* if it is not *inconsistent*.

Definition 2.10 (MCS). A *maximal consistent set* is a set of formulas which is consistent and of which any proper superset is inconsistent.

It is easy to prove, using the appropriate variation on Lindenbaum's lemma, see for example [3], that any formula consistent with $\Sigma\mathrm{L}$ is contained in some maximal consistent set. Also, for any model we can associate to each world an MCS, namely the set of formulas true in that world. The road followed in most modal completeness proofs is: Fix a nonprovable formula A and some MCS Δ_A containing $\neg A$ and find a model which contains a world w such that the set of formulas true at w is Δ_A.

In general MCS's are infinite. If there are infinitely many formulas of the form $\Diamond B$ in an MCS then we might need infinitely many successors.

So in cases in which one wants a finite model, and this is such a case, one needs to refine a bit.[5]

One solution is to take maximal sets which are maximal as a subset of some (fixed) finite set.

Another is to take (full) MCS's but only to require the model to contain a world w such that a finite number of formulas in Δ_A are true at w. We take the latter option. What finite subset of Δ_A to take depends on the particular nonprovable formula. So we make the following definition:

Definition 2.11 (Relevant Set). If A is a formula then define $\mathcal{R}_A$, *the relevant set of* A, to be the smallest set such that

1. $\neg A \in \mathcal{R}_A$,
2. $\mathcal{R}_A$ is closed under subformulas and single negation.

Next let us have some definitions to talk about MCS's in combination with frames.

Definition 2.12 (Labeled frames). A quadruple $\langle W, R, S, \nu \rangle$ is a *labeled frame* if

- $\langle W, R, S \rangle$ is a frame and
- ν is a function $W \longrightarrow \{x \mid x \text{ an MCS}\}$.

Definition 2.13 ($\prec$, $\subseteq_{\Sigma_1}$). Let Δ_0, Δ_1 and Γ be MCS's. Define the binary relations $\prec$ and $\subseteq_{\Sigma_1,\Gamma}$ as follows:

- $\Delta_0 \prec \Delta_1 \Leftrightarrow \{D, \Box D \mid \Box D \in \Delta_0\} \subseteq \Delta_1$,
- $\Delta_0 \subseteq_{\Sigma_1,\Gamma} \Delta_1 \Leftrightarrow \{D \in \Delta_0 \mid \Sigma_1 D \in \Gamma\} \subseteq \Delta_1$.

[5] In any case, GL is not compact (see [2]) so the use of infinite MCS's in this way is bound to fail anyway.

How will we be using labeled frames? We fix a nonprovable formula A, an MCS Δ_A which contains $\neg A$ and construct a labeled frame F with a world w such that $\nu^F(w) = \Delta_A$. We can define a model M from a labeled frame F by putting $V^M(p) = \{w \mid p \in \nu^F(w)\}$.

As mentioned above we want a finite set of the formulas in $\Delta_A (= \nu^F(w))$ (namely the set $\mathcal{R}_A \cap \Delta_A$) to be formulas forced at $w \in M$, in fact we will construct F in such a way that for each v in F the formulas $\mathcal{R}_A \cap \nu^F(v)$ are true at v. Let us say that 'a truth lemma holds' if the frame at hand possesses this property. This is somewhat imprecise since we do not specify the formula A but it will be clear what is meant. Now labeled frames could possess certain 'problems' which prevent such a truth lemma to hold:

Definition 2.14 (X-problems). Suppose X is a set of formulas, F a labeled frame. An X-*problem* in F is a node $w \in F$ and a formula $A \in X \cap \nu^F(w)$ such that one of the following two cases applies:

1. $A = \neg\Sigma_1 B$ and for no $uS^F_w v$: $B \in \nu^F(u)$ and $\neg B \in \nu^F(v)$,
2. $A = \neg\Box B$ and for no $wR^F v$: $\neg B \in \nu^F(v)$.

How are we to determine an F without any such problems? We take the limit of a series of better and better approximations: F will be the union of a chain $F_0 \subseteq F_1 \subseteq \cdots$ of labeled frames where each frame F_{i+1} has fewer 'problems' than its predecessor F_i. In view of this goal the usefulness of the following lemma is evident.

Lemma 2.15. *If F and G are two labeled frames such that $F \subseteq G$. Then if $w \in F$ and (w, A) is an X-problem in G then (w, A) is an X-problem in F as well.*

Proof. Trivial $\dashv$

Besides having no problems we need (for a truth lemma to hold) that F is reasonable:

Definition 2.16 (Reasonable). We say that a labeled frame F is *reasonable* if:

- $vR^F w \Rightarrow \nu^F(v) \prec \nu^F(w)$,
- $vS^F_u w \Rightarrow \nu^F(v) \subseteq_{\Sigma_1, \nu^F(u)} \nu^F(w)$.

Finally, the resulting frame should satisfy all the ΣL-frame properties for a truth lemma to imply the completeness theorem. It is easy to see that the intersection of any set of (labeled) frames that satisfy all the ΣL properties satisfies the ΣL properties as well. If there is at least one frame which satisfies all the ΣL-frame properties extending a given one then we can talk about the smallest extension of a frame that satisfies all the ΣL properties.

Definition 2.17 (ΣL-closure). If $F = \langle W, R, S\rangle$ is a finite frame such that R is conversely well-founded. Then we define the ΣL-*closure* of F to be the intersection of all the ΣL-frames with domain W (the same domain as F) and which extends F. If F is a labeled frame then the ΣL-closure of F is $\langle W^G, R^G, S^G, \nu^F\rangle$ where G is the ΣL closure of $\langle W^F, R^F, S^F\rangle$.

Lemma 2.18. *Suppose $F = \langle W, R, S, \nu\rangle$ is a finite labeled frame such that R is conversely well founded. If F is reasonable then the* ΣL-*closure of F is reasonable as well.*

Proof. It is easy to see that, since the ΣL-closure of a frame F is a ΣL-frame. And is the smallest ΣL-frame extending F. The ΣL closure of a frame F is equal to the union of a chain $F_0 \subseteq F_1 \subseteq \cdots$ which satisfies[6]:

1. $F_0 = F$ and
2. if $i \geq 0$ and $F_i = \langle W, R, S, \nu\rangle$ one of the following applies:
    (a) $aRbRc$ but not aRc and $F_{i+1} = \langle W, R + \{(a, c)\}, S, \nu\rangle$,
    (b) aS_wbRc but not aRc and $F_{i+1} = \langle W, R + \{(a, c)\}, S, \nu\rangle$,
    (c) $wRaRb$ but not aS_wb and $F_{i+1} = \langle W, R, S + \{(w, a, b)\}, \nu\rangle$,
    (d) wRv and aS_vb but not aS_wb and $F_{i+1} = \langle W, R, S + \{(w, a, b)\}, \nu\rangle$,
    (e) wS_tv and aS_vb but not aS_wb and $F_{i+1} = \langle W, R, S + \{(w, a, b)\}, \nu\rangle$,
    (f) F_i is a ΣL-frame and $F_{i+1} = F_i$.

We show with induction on i that each F_i is reasonable. The case $i = 0$ is trivial. So assume that $i \geq 0$ and F_i is reasonable. Assume we are in Case 2e. Then $R^{F_i} = R^{F_{i+1}}$. Moreover for all i: $W^{F_i} = W$ and $\nu^{F_i} = \nu$. So

$$F_i = \langle W, R, S^{F_i}, \nu\rangle,$$
$$F_{i+1} = \langle W, R, S^{F_{i+1}}, \nu\rangle.$$

We thus have to show:

$$\text{for all } a, b\text{: } aRb \Rightarrow \nu(a) \prec \nu(b), \tag{3}$$
$$\text{for all } a, b, w\text{: } aS_w^{F_{i+1}}b \Rightarrow \nu(a) \subseteq_{\Sigma_1,\nu(w)} \nu(b). \tag{4}$$

By (IH) (3) holds directly. So what is left is to show (4). Pick a, b, w and assume $aS_w^{F_{i+1}}b$, $D \in \nu(a)$ and $\Sigma_1(D) \in \nu(w)$. If $aS_w^{F_i}b$ then we are done since F_i is reasonable so we can assume that for some u, t: $wS_t^{F_i}u$ and $aS_u^{F_i}b$. By (IH):

$$\nu(w) \subseteq_{\Sigma_1,\nu(t)} \nu(v), \tag{5}$$
$$\nu(a) \subseteq_{\Sigma_1,\nu(v)} \nu(b). \tag{6}$$

Since $\Sigma_1\Sigma_1 D \in \nu(t)$ we have by (5): $\Sigma_1 D \in \nu(v)$. And thus by (6): $D \in \nu(b)$. Which shows (4). The other cases are even easier.

$\dashv$

[6]Since F is finite any such chain will do. If we allow F to be infinite we should be more careful.

2.3.2 Tools

Lemmas 2.20 and 2.21 are the main engine behind the definition of the chain of approximations. Before we can prove them we need some well-known facts about GL. For proofs see for example [2][15].

Lemma 2.19. *Suppose X is a finite set of formulas. Then*

1. $\Sigma\mathrm{L} \vdash \bigwedge \boxdot X \to (\Box A \to A) \Rightarrow \Sigma\mathrm{L} \vdash \bigwedge \boxdot X \to A$.
2. *If $X + \{A\}$ is $\Sigma\mathrm{L}$-consistent. Then $\boxdot X + \{A, \Box\neg A\}$ is $\Sigma\mathrm{L}$-consistent as well.*

$\dashv$

Lemma 2.20. *Let w be an* MCS. *Suppose $\neg\Sigma_1 A \in w$. Then there exist u, v s.t. $w \prec u, v$, $A \in u$, $\neg A \in v$ and $u \subseteq_{\Sigma_1, \nu(w)} v$. Moreover we can ensure that $\Sigma_1 A \in u, v$.*

Proof. Let

$$\begin{array}{rcl} \boxdot(x) & = & \{D, \Box D \mid \Box D \in x\}, \\ \Sigma(x) & = & \{D \mid \Sigma_1 D \in x\}, \\ \Sigma_{con} & = & \{Y \subseteq \Sigma(w) \mid \{\neg A\} + \boxdot(w) + Y \text{ is consistent and maximally such}\}. \end{array}$$

Although we do not strictly need to show it separately, let us first prove that Σ_{con} is not empty. The argument is a simplification of the complete proof of Lemma 2.20.

For Σ_{con} to be not empty it is sufficient that $\{\neg A\} + \boxdot(w)$ is consistent. Suppose, for a contradiction, that this is not the case. Then for some finite $w' \subseteq w$:

$$\vdash \bigwedge \boxdot(w') \to A,$$

thus

$$\vdash \Box(\bigwedge \boxdot(w')) \to \Box A,$$

but trivially $\vdash \Box A \to \Sigma_1 A$ so

$$\Sigma_1 A \in w.$$

A contradiction.

If we write $\vdash \bigwedge \boxdot(w) \to A$ above as $\vdash \bigwedge \boxdot(w) \wedge \neg A \to \bot$, then the argument below is an extension of the argument above replacing $\bot$ by a more complex Σ_1-sentence.

Claim. For some $Y \in \Sigma_{con}$ the set

$$\{A, \Sigma_1 A\} + \boxdot(w) + \{\neg\sigma \mid \sigma \in \Sigma(w) - Y\}$$

is consistent.

Proof of Claim. Suppose the claim is false. Then we can choose for each $Y \in \Sigma_{con}$ a finite set $F_Y \subseteq \Sigma(w) - Y$ such that

$$\{A, \Sigma_1 A\} + \boxdot(w) + \{\neg\sigma \mid \sigma \in F_Y\} \tag{7}$$

is inconsistent. Next we will show that:

$$\{\neg A\} + \boxdot(w) + \{\bigvee_{\sigma \in F_Y} \sigma \mid Y \in \Sigma_{con}\} \text{ is inconsistent.} \tag{8}$$

For suppose (8) is not the case. Then for some $S \in \Sigma_{con}$ (note that $\Sigma(w)$ is closed under disjunctions):

$$\{\bigvee_{\sigma \in F_Y} \sigma \mid Y \in \Sigma_{con}\} \subseteq S.$$

In particular we have

$$\bigvee_{\sigma \in F_S} \sigma \in S.$$

But for all $\sigma \in F_S$ we have

$$\sigma \notin S,$$

in contradiction with the maximality of S. Thus we have shown (8). So we can select some finite $Y' \subseteq \Sigma_{con}$ and a finite $w' \subseteq w$ such that

$$\vdash \neg A \wedge \bigwedge \boxdot(w') \rightarrow \neg \bigwedge_{Y \in Y'} \bigvee_{\sigma \in F_Y} \sigma \tag{9}$$

By the inconsistency of the sets (7) for each $Y \in \Sigma_{con}$ there exists a finite $w_Y \subseteq w$ such that

$$\vdash A \wedge \Sigma_1 A \wedge \bigwedge \boxdot(w_Y) \rightarrow \neg \bigwedge_{\sigma \in F_Y} \neg\sigma.$$

So we certainly have

$$\vdash A \wedge \Sigma_1 A \wedge \bigwedge \boxdot(\bigcup_{Y \in Y'} w_Y) \rightarrow \neg \bigvee_{Y \in Y'} \bigwedge_{\sigma \in F_Y} \neg\sigma. \tag{10}$$

Combining (9) with (10) we get

$$\vdash \Box \bigwedge \boxdot(w' \cup \bigcup_{Y \in Y'} w_Y) \wedge \Box\Sigma_1 A \rightarrow \Box(A \leftrightarrow \bigwedge_{Y \in Y'} \bigvee_{\sigma \in F_Y} \sigma).$$

Thus

$$\vdash \Box \bigwedge \boxdot(w' \cup \bigcup_{Y \in Y'} w_Y) \rightarrow (\Box\Sigma_1 A \rightarrow \Sigma_1 A).$$

And by Lemma 2.19

$$\vdash \bigwedge w' \cup \bigcup_{Y \in Y'} w_Y \rightarrow \Sigma_1 A.$$

In contradiction with $\neg\Sigma_1 A \in w$, and thus the claim must be true. $\dashv$

So, to summarize, we have for some $Y \in \Sigma_{con}$ that both the sets

$$\{A, \Sigma_1 A\} + \boxdot(w) + \{\neg\sigma \mid \sigma \in \Sigma(w) - Y\} \tag{11}$$

$$\{\neg A\} + \boxdot(w) + Y \tag{12}$$

are consistent. Since $\Sigma_1 A$ then must be in Y the lemma follows by taking u, v extending (11) and (12) respectively. $\dashv$

Lemma 2.21. *Suppose w is an* MCS. *Suppose $\neg\Box D \in w$. Then there exists v such that $w \prec v$ and $\neg D \in v$. Moreover we can choose v such that $\neg\Box D \notin v$.*

Proof. The usual proof already gives a v s.t. $\Box D \in v$. See for example [2][10]. $\dashv$

Now we put the previous two lemmas to work in showing that once we have a reasonable frame F_i we can extend it to a reasonable frame F_{i+1} with fewer problems. Before we state and prove this theorem we need to measure the number of X-problems in a node $v \in F_i$. We will do this for finite X only and we will not count them exactly but bound them from above only.

Definition 2.22 ($|v|_X$). If $F = \langle W, R, S, \nu\rangle$ is a labeled frame, $v \in W$ and X a finite set of formulas then define:

$$|v|_X = \#\{\neg\Sigma_1 A \in \nu(v) \cap X\} + \#\{\neg\Box A \in \nu(v) \cap X\}.$$

2.3.3 Extension theorem

Theorem 2.23 (Extension theorem). *Let F be a labeled frame. If F is reasonable and (w, A) an X-problem in F then, there exists a labeled frame G such that: $F \subseteq G$, G is reasonable and (w, A) is not an X-problem in G. Moreover if X is finite then:*

$$v \in G - F \Rightarrow |v|_X < |w|_X.$$

And if F is finite then so is G.

Proof. We treat the case $A = \neg\Sigma_1 B$. The case $A = \neg\Box B$ goes similarly. Let u, v be two nodes not in F. Δ_0, Δ_1 be two MCS's such that $\nu(w) \prec \Delta_0, \Delta_1$, $\Delta_0 \subseteq_{\Sigma_1, \nu^F(w)} \Delta_1$, $\Sigma_1 B, B \in \Delta_0$ and $\Sigma_1 B, \neg B \in \Delta_1$. These Δ_0, Δ_1 exist by Lemma 2.20. Now put

$$G = \langle W^F + \{u, v\}, R^F + \{(w, u), (w, v)\}, S^F + \{(w, u, v)\}, \nu^F + \{(u, \Delta_0), (v, \Delta_1)\}\rangle.$$

Clearly G is reasonable and (w, A) is not an X-problem in G. Now suppose X is finite. Since for each C: $\Sigma\mathrm{L}\vdash \Box C \to \Box\Box C$, $\Sigma\mathrm{L}\vdash \Sigma_1 C \to \Box\Sigma_1 C$ and $\nu(w) \prec \Delta_0, \Delta_1$ we have $|u|_X \leq |w|_X$ and $|v|_X \leq |w|_X$. So since $\Sigma_1 B \in \nu(u) - \nu(w)$ and $\Sigma_1 B \in \nu(v) - \nu(w)$, we conclude: $|u|_X < |w|_X$ and $|v|_X < |w|_X$. $\dashv$

Proof of Theorem 2.8. Let A be some formula not provable in ΣL, let Δ_A be some MCS containing $\neg A$ and let $\mathcal{R}_A$ be the relevant set of A. We define a chain $F_0 \subseteq F_1 \subseteq \cdots$ of labeled frames inductively.

- $F_0 = \langle \{w_0\}, \emptyset, \emptyset, \{(w_0, \Delta_A)\}\rangle$,
- If F_i is defined then
  - If F_i is not a ΣL-frame then let F_{i+1} be the ΣL-closure of F_i,
  - Else if F_i contains no $\mathcal{R}_A$-problems let $F_{i+1} = F_i$,
  - Else let (x, B) be some $\mathcal{R}_A$-problem and apply the Extension theorem (with $\mathcal{R}_A$ for X and (x, B) for (w, A)) to find F_{i+1}.

It is evident that each F_i is reasonable and finite. We show that $F = \bigcup_{i\geq 0} F_i$ is a finite ΣL-frame without any $\mathcal{R}_A$-problems.

If we combine Lemma 2.15 (once we have solved a problem it will not reoccur) with the fact that $\mathcal{R}_A$ is finite (for each node v there are only finitely many problems involving v) we see that F is finitely branching: for each $w \in F$, $\{v \mid wR^F v\}$ is finite. In addition we have $wR^F v \Rightarrow |w|_{\mathcal{R}_A} > |v|_{\mathcal{R}_A} \geq 0$.

Conclusion: F is finite and thus for some j we have $j' \geq j \Rightarrow F_{j'} = F_j$. This latter fact implies that F is a ΣL-frame and does not have any $\mathcal{R}_A$-problems.

If we define $M = \langle W^F, R^F, S^F, V\rangle$ by putting $V(p) = \{w \in W^F \mid p \in \nu^F(w)\}$ then one proves by induction on B that for each $w \in M$ and $B \in \nu^F(w) \cap \mathcal{R}_A$: $w \models B$. So in particular $w_0 \models \neg A$. ⊣

2.4 Failure of interpolation

Definition 2.24 (Interpolant). Suppose $\Sigma\mathrm{L} \vdash A \to B$. We say that I is an *interpolant* for $A \to B$ if all the proposition variables in I occur in both A and B and $\Sigma\mathrm{L} \vdash A \to I$ and $\Sigma\mathrm{L} \vdash I \to B$.

Alternatively one can say that if $\{A, B\}$ is inconsistent then I is an interpolant for $\{A, B\}$ if all proposition variables in I occur in both A and B and $\Sigma\mathrm{L} \vdash A \to I$ and $\Sigma\mathrm{L} \vdash B \to \neg I$.

Theorem 2.25. *If $M = \langle W, R, S, V\rangle$ is a ΣL-model and $\overline{M} = \langle W, R, \overline{S}, V\rangle$, where $\overline{S}$ is the unique ternary relation on W such that for each $w \in W$: $\overline{S}_w$ is the transitive closure of S_w. Then*

1. *$\overline{M}$ is a ΣL-model and*
2. *for each formula I and each $x \in W$: $\overline{M}, x \models I \Leftrightarrow M, x \models I$.*

Proof. 1 An easy verification of the properties 2.5.

2 Induction on the complexity of I. Boolean connectives and $\Box$ cases are trivial, so suppose $I = \Sigma I_0$. ($\Leftarrow$) Take x_0, x_1 s.t. xRx_0, x_1 and for some xRw $x_0\overline{S}_w x_1$ $x_0 \models I_0$. Take any sequence $x_0 = z_0 S_w z_1 S_w \cdots S_w z_k = x_1$. $z_1 \models I_0$ and thus $z_2 \models I_0$ and thus … and thus $z_k = x_1 \models I$. Which was to be proved. ($\Rightarrow$) Trivial.

⊣

Definition 2.26 (ΣL-Bisimulation). Let M_0 and M_1 be ΣL-models and P a set of proposition variables. A binary relation $Z \subseteq M_0 \times M_1$ is a *ΣL-Bisimulation* w.r.t. P if the following conditions are met.

1. If $a_0 Z a_1$ then for each $p \in P$: $a_0 \in V_0(p) \Leftrightarrow a_1 \in V_1(p)$.
2. If $a_0 Z a_1$ and $a_0 R b_0$ then there exists b_1 such that $b_0 Z b_1$ and $a_1 R b_1$.
3. Same as 2 with M_0 and M_1 interchanged.
4. If $a_0 Z a_1$ then for all $b_0, c_0 \in M_0$. If
$$a_0 R b_0, c_0 \text{ and } b_0 S_{a_0} c_0,$$
then there exist $b_1, c_1 \in M_1$ such that $b_0 Z b_1$ and $c_0 Z c_1$
$$a_1 R b_1, c_1 \text{ and } b_1 S_{a_1} c_1.$$
5. Same as 4 with M_0 and M_1 interchanged.

Notationally this notion of bisimulation is somewhat complex. However, in essence, it is in fact simpler than the notion of bisimulation for ILM (Definition 2.30 below). If you look at satisfaction as a game, then falsifying a formula ΣA in a point x requires Falsifier to simultaneously pick xRy, z with $yS_x z$ whereas in ILM the falsification of a formula $A \rhd B$ requires two moves. First Falsifier picks $xRy \models A$ and then Verifier should (not be able to) pick some $xRz \models B$ with $yS_x z$.

Theorem 2.27. *If M_0 and M_1 are ΣL-models and Z is a ΣL-bisimulation w.r.t. P between them, then for any formula I whose proposition variables all occur in P, $m_0 \in M_0$, $m_1 \in M_1$ and $m_0 Z m_1$ then*

$$M_0, m_0 \models I \Leftrightarrow M_1, m_1 \models I$$

Proof. Induction on I. Let $m_0 \in M_0$ and $m_1 \in M_1$ be such that $m_0 Z m_1$. The atomic and boolean connective cases are trivial.

Case $I = \Box I_0$. ($\Leftarrow$) Assume $M_0, m_0 \not\models \Box I_0$. Then there exists n_0 such that $m_0 R n_0$ and $M_0, n_0 \not\models I_0$. Then by 2 of Definition 2.26 there exists n_1 such that $m_1 R n_1$ and $n_0 Z n_1$. But then by (IH), $M_1, n_1 \not\models I_0$. ($\Rightarrow$) Analogously.

Case $I = \Sigma_1 I_0$. ($\Leftarrow$) Assume $M_0, m_0 \not\models \Sigma_1 I_0$. Then there exist n_0, l_0 such that $m_0 R n_0, l_0$, $n_0 S_{m_0} l_0$ and $M_0, n_0 \models I_0$ and $M_0, l_0 \not\models I_0$. By 4 of Definition 2.26 there exist n_1, l_1 such that $m_1 R n_1, l_1$ $n_1 S_{m_1} l_1$ and $n_0 Z n_1$ and $l_0 Z l_1$. Then (IH) $M_1, n_1 \models I_1$ and $M_1, l_1 \not\models I_1$. ($\Rightarrow$) analogous $\dashv$

Immediately the question arrises whether Theorem 2.27 can be reversed. In Corollary 2.29 below this is answered negatively.

Clearly definition 2.24 can be given for all sorts of modal logics (e.g. GL, ILM, K4 etc.). We say that such a logic (Lgc say) has the interpolation property, or simply has interpolation, whenever Lgc $\vdash A \to B$ implies that $A \to B$ has an interpolant. Now we formulate the main result of this section:

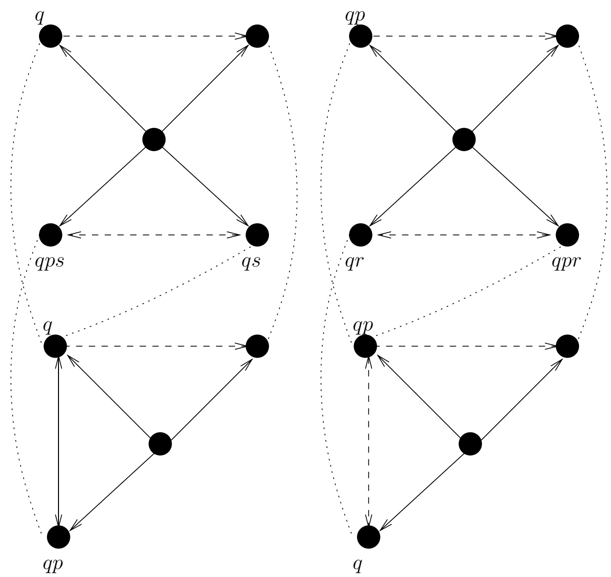


Figure 1: Two bisimularions.

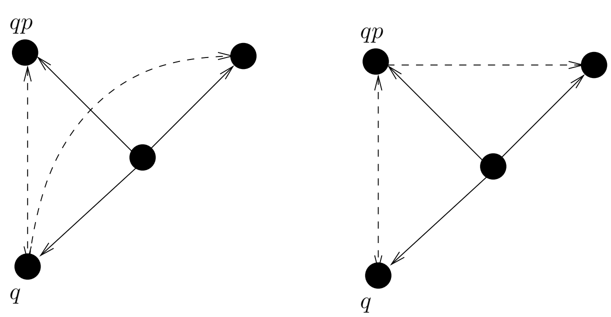


Figure 2: Almost the same models.

Theorem 2.28. *The logic* ΣL *does not have interpolation.*

Proof. Consider the following two formulas

$$\begin{aligned} A(p,q,s) &=_{\rm def} \quad \neg\Sigma_1 q \wedge \Sigma_1 s \wedge \Box(s \to q) \wedge \Box(p \wedge q \to s), \\ B(p,q,r) &=_{\rm def} \quad \neg\Sigma_1 q \wedge \Sigma_1 r \wedge \Box(r \to q) \wedge \Box(\neg p \wedge q \to r). \end{aligned}$$

Write $A_s = A(p,q,s)$ and $B_r = B(p,q,r)$. Then $\{A_s, B_r\}$ is inconsistent but for no formula $I = I(p,q)$ with all proposition variables among p, q we have $\Sigma\mathrm{L} \vdash A_s \to I$ and $\Sigma\mathrm{L} \vdash B_r \to \neg I$. First their inconsistency. We have

$$\vdash A_s \wedge B_r \to \Box(q \leftrightarrow r \vee s),$$

as well as

$$\vdash A_s \wedge B_r \to \Sigma_1(r \vee s).$$

Thus

$$\vdash A_s \wedge B_r \to \Sigma_1 q.$$

So indeed $\{A_s, B_r\}$ is inconsistent.

Now, to show that no interpolant exists, consider Figure 1. The upper left model forces A_s in its center point and the upper right model forces B_r in its center point. The upper left model is bisimilar w.r.t. $\{p,q\}$ with the lower left as indicated by the dotted lines. Similarly for the two models on the right. Moreover if we replace in the lower two models the S_w relations by their transitive closures, the two models become the same. This is clarified by Figure 2, where the lower two models of Figure 1 are drawn again but the left one is layed out a bit differently. Applying Theorem 2.25 and Theorem 2.27 we see that the upper two models force the same $\{p,q\}$-formulas in their center point. So clearly an interpolant for $A_s \to \neg B_r$ cannot exist. $\dashv$

From the above proof we can immediately extract the following corollary.

Corollary 2.29. *There exist two models* M *and* M' *and two worlds* $m \in M$ *and* $m' \in M'$ *such that* m *and* m' *force the same formula. But there does not exist a bisimulation between* M *and* M' *that connects* m *and* m'*.*

Proof. As was shown, the two center-points of (the restrictions to the proposition variables $\{p.q\}$ of) the two upper models of Figure 1 force the same formulas. One easily verifies directly that there does not exists a ($\{p,q\}$-)bisimulation that connects those two points[7]. $\dashv$

Two questions arise.

1. What additions to the language would make an interpolant of $A_s \to \neg B_r$ exist and

[7]Below one can find an indirect argument for this fact, namely that no two bisimilar model can distinguish between A_s and B_r.

2. Can we simplify the proof by using two bisimilar models (the models in Figure 2 are not bisimilar).

To answer the first question we use the idea of an arithmetical semantics for a modal logic. The reader should first read Section 2.5 if (s)he is unfamiliar with this. A sufficient addition for an interpolant is Σ_1-interpolability[7]:

$$I_{\Sigma_1}(\phi, \psi) = \exists \sigma (\Sigma_1(\sigma) \wedge \Box(\phi \to \sigma) \wedge \Box(\sigma \to \psi).$$

For any mapping $* : \mathrm{PROP} \longrightarrow$ 'arithmetical sentences'[8] clearly:

$$\mathrm{PA} \vdash A(p^*, q^*, s^*) \to I_{\Sigma_1}(p^* \wedge q^*, q^*) \tag{13}$$

and

$$\mathrm{PA} \vdash B(p^*, q^*, r^*) \to I_{\Sigma_1}(\neg p^* \wedge q^*, q^*).$$

Moreover

$$\mathrm{PA} \vdash \Sigma_1 q^* \leftrightarrow I_{\Sigma_1}(q^*, q^*).$$

Lemma. *For all* ϕ_0, ϕ_1 *and* ψ. $\mathrm{PA} \vdash I_{\Sigma_1}(\phi_0, \psi) \wedge I_{\Sigma_1}(\phi_1, \psi) \to I_{\Sigma_1}(\phi_0 \vee \phi_1, \psi)$.

Proof. Reason in PA. Assume that for some σ_0: $\Sigma_1(\sigma_0)$, $\mathrm{PA} \vdash \phi_0 \to \sigma_0$ and $\mathrm{PA} \vdash \sigma_0 \to \psi$. And assume that for some σ_1: $\Sigma_1(\sigma_1)$, $\mathrm{PA} \vdash \phi_1 \to \sigma_1$ and $\mathrm{PA} \vdash \sigma_1 \to \psi$. Then $\mathrm{PA} \vdash \sigma_0 \vee \sigma_1 \to \psi$ and $\mathrm{PA} \vdash \phi_0 \vee \phi_1 \to \sigma_0 \vee \sigma_1$. Since $\Sigma_1(\sigma_0 \vee \sigma_1)$, this concludes the proof. ⊣

So, by the above lemma:

$$\mathrm{PA} \vdash I_{\Sigma_1}(\neg p^* \wedge q^*, q^*) \wedge I_{\Sigma_1}(p^* \wedge q^*, q^*) \to I_{\Sigma_1}(q^*, q^*).$$

So

$$\mathrm{PA} \vdash \neg \Sigma_1 q^* \wedge I_{\Sigma_1}(\neg p^* \wedge q^*, q^*) \to \neg I_{\Sigma_1}(p^* \wedge q^*, q^*).$$

And thus

$$\mathrm{PA} \vdash B(p^*, q^*, r^*) \to \neg I_{\Sigma_1}(p^* \wedge q^*, q^*). \tag{14}$$

So, combining (13) and (14), if we could express Σ_1-interpolability, by a (modal) formula $I_{\Sigma_1}(p, q)$ say, then the formula $I_{\Sigma_1}(p \wedge q, q)$ is an interpolant for $A_s \to \neg B_r$.

In [7] a modal theory, which language contains, besides the $\Box$, an operator for Σ_1-interpolability, is developed. And indeed this logic has the interpolation property. This logic is also evaluated on Veltman frames and the same notion of bisimulation that we used is appropriate for this extended[9] logic. But this answers the second question: We will not be able to give two bisimilar models that differentiate between A_s and B_r. (Since in this enriched language of Σ_1-interpolabilty we do have an interpolant.) The harmless looking Theorem 2.25 is thus an essential part of our present argument that shows Theorem 2.28.

[8] Actually $*$ should map to Gödel numbers of arithmetical sentences but we are only sketching an idea.

[9] Σ_1 is definable in the language of Σ_1-interpolability.

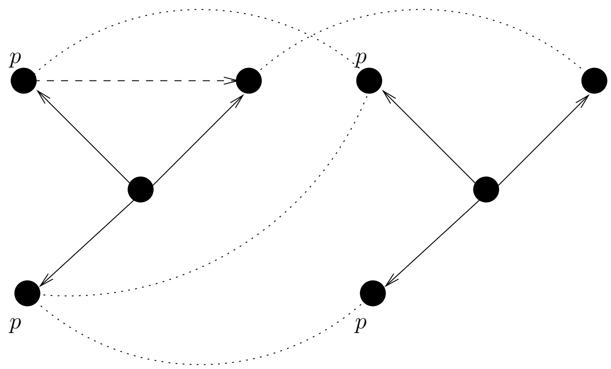


Figure 3: $\rhd$-bisimilar but not σ-bisimilar.

With one of the goals of our investigation in mind: extending ILM in such a way that we get interpolation, it is reasonable to think about an extension of the present result to a combination of ILM with ΣL. Let us therefore state and investigate a notion of bisimulation for IL-formulas.

Definition 2.30 (IL-bisimulation). Let M and M' be two models and P a set of proposition letters. A relation $Z \subseteq M \times M'$ is an *IL-bisimulation w.r.t.* P if

1. $wZw' \Rightarrow$ for each $p \in P$: $w \in V(p) \Leftrightarrow w' \in W(p)$,
2. If wZw' and wRv then there exists some vZv' such that $w'Rv'$ and for each $v'S_{w'}u'$ there exists some uZu' with vS_wu,
3. Same as 2 with M and M' interchanged.

Of course we have a theorem asserting that two IL-bisimilar (w.r.t. P) models force the same IL-formulas (with proposition letters in P) in bisimilar points [18].

One difficulty arises. The two notions of bisimulation are incomparable. We say that two models are IL(ΣL)-bisimilar if a IL(ΣL)-bisimulation w.r.t. PROP exists.

Fact 2.31. *There are two models which are* ΣL*-bisimilar but not* IL*-bisimilar. And there are two models which are* IL*-bisimilar but not* ΣL*-bisimilar.*

Proof. The proof is contained in Figures 3 and 4. Namely the claimed bisimulations are indicated and the models in 4 are distinguished by $p \rhd \neg p$ and those in Figure 3 by $\Sigma_1 p$. $\dashv$

A notion of bisimulation which handles both types of formulas is thus strictly stronger than the two separate notions. In Section 3.6 below we will see that it is sufficient to use only these two separate notions of bisimulation in order to extend the result of this section to a combined logic of ΣL and ILM.

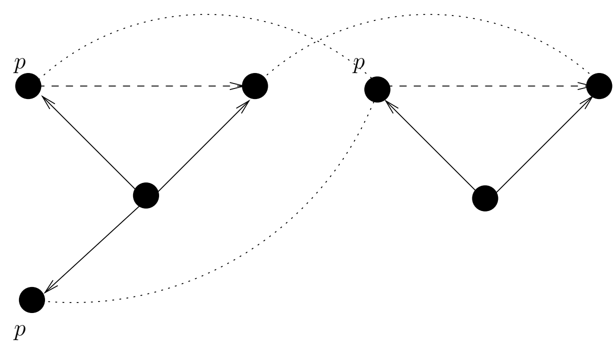


Figure 4: σ-bisimilar but not $\rhd$-bisimilar.

2.5 Arithmetical interpretation

It is possible to give an arithmetical meaning to ΣL-formulas. We will however first extend the logic ΣL to a logic ΣILM, give arithmetical meaning to the ΣILM-formulas and project this on ΣL. The reader is referred to Section 4.2 for details.

3 The logic ΣILM

In Section 3.1 we elaborate some more on ILM. In Sections 3.2 and 3.3 the logic ΣILM and a modal semantics is introduced. This semantics are not the most obvious one. In Section 3.3.1 we explain why we deviate from the more obvious choice. In Section 3.4 and 3.5 we show that ΣILM is sound and complete w.r.t. the proposed semantics. In Section 3.6 we extend the results in Section 2.4 to show that ΣILM lacks interpolation. In Section 4 we give arithmetical meaning to ΣILM and in Section 4.2 we reflect this on ΣL.

3.1 ILM

The interpretability logic of PA known as ILM has been introduced in Section 1.1. To investigate this logic we can evaluate IL-formulas on the same frames we have used for ΣL. But first let us prove a lemma.

Lemma 3.1. *The following is provable in* IL.

1. $\Box\neg A \leftrightarrow A \rhd \bot$,
2. $A \vee \Diamond A \rhd A$.
3. $A \wedge \Box\neg A \rhd B \rightarrow A \rhd B$,

Proof. (1.) By IL Axiom 6: IL $\vdash A \rhd \bot \rightarrow (\neg\Diamond\bot \rightarrow \neg\Diamond A)$. Trivially IL $\vdash \neg\Diamond\bot$, so IL $\vdash A \rhd \bot \rightarrow \Box\neg A$. The other direction is simply an instantiation of Axiom

3. (2.) By Axiom 3: $\mathrm{IL} \vdash A \rhd A$. 2 follows if we combine this via Axiom 5 with Axiom 7. (3.) We use the following fact about GL (see [10]):

$$\mathrm{GL} \vdash A \to (A \wedge \Box\neg A) \vee \Diamond(A \wedge \Box\neg A).$$

Using necessitation and Axiom 3: $\mathrm{IL} \vdash A \rhd (A \wedge \Box\neg A) \vee \Diamond(A \wedge \Box\neg A)$ and then using part 2 of this lemma and Axiom 4:

$$\mathrm{IL} \vdash A \rhd A \wedge \Box\neg A.$$

So, again using Axiom 4, we conclude $\mathrm{IL} \vdash A \wedge \Box\neg A \rhd B \to A \rhd B$. ⊣

Note that the first item allows us to only consider $\rhd$ and define $\Box A$ as $(\neg A) \rhd \bot$.

Definition 3.2 (IL forcing relation). If $M = \langle W, R, S, V\rangle$ such that $\langle W, R, S\rangle$ is a Veltman frame and V is a valuation on W then let $\models_M$ be *the unique relation between worlds $w \in M$ and IL-formulas* which satisfies:

1. $w \models_M p \Leftrightarrow w \in V^M(p)$,
2. $w \models_M A \rhd B \Leftrightarrow$ for each wRu such that $u \models_M A$ there exists uS_wv such that $v \models_M B$,
3. The usual clauses for propositional connectives.

The proper Veltman frames differ slightly from those for ΣL.

Definition 3.3 (IL-frame). A triple $F = \langle W, R, S\rangle$ is called an *IL-frame* if F is a frame and

1. R is transitive and conversely well-founded,
2. for all a, b, c, w, t
   (a) $wRaRb \Rightarrow aS_wb$,
   (b) $aS_wbS_wc \Rightarrow aS_wc$,
   (c) $wRa \Rightarrow aS_wa$.

Definition 3.4 (ILM-frame). A triple $F = \langle W, R, S\rangle$ is an *ILM-frame* if it is an IL-frame and satisfies the additional requirement (M):

$$aS_wbRc \Rightarrow aRc.$$

As usual we can talk about the validity of an IL-formula on a frame and it turns out that we have the following theorem (see [10][4]).

Theorem 3.5. *The logic* IL(M) *is sound and complete w.r.t. finite* IL(M)-*frames.*

3.2 The logic ΣILM

In this section we begin the development of an arithmetically complete logic which can talk about both Π_1-Conservativity and Σ_1-ness. The Π_1-Conservativity part is taken to be ILM except for the (M) axiom, which we can replace by $\Sigma_1 C \wedge (A \rhd B) \to A \wedge C \rhd B \wedge C$. The Σ_1 part of the logic is simply ΣL from Section 2. The resulting logic will we denote by ΣILM.

Definition 3.6 (ΣILM-formulas). *ΣILM-formulas* are built up from variables in PROP, propositional connectives, unary operators $\Box$ and Σ_1 and the binary operator $\rhd$.

Definition 3.7 (ΣILM). ΣILM is the set of schemata IL + ΣL together with the schema $M(\Sigma)$:

$$\Sigma_1 C \wedge (A \rhd B) \to A \wedge C \rhd B \wedge C.$$

The logic ΣILM is the smallest set of ΣILM-formulas which contains the universal closure of all the ΣILM schemata and is closed under modus ponens and necessitation. We write $\Sigma ILM \vdash A$, or $\vdash A$, for $A \in$ *the logic* ΣILM.

Ideally we would like to prove this logic to be complete with respect to the class of frames which both satisfy the ΣL- and ILM-frame conditions. And this can indeed be done except for the fact that we must then allow for infinite models.

This, however, makes the situation somewhat problematic if we want to show the logic to be arithmetically complete. As in [1] we could introduce the notion of a primitive recursive model. These are models which, among other things, have as their domain a primitive recursive set of numbers so that (for instance) PA can easily talk about them, almost as easily as with finite models (which are a particular case of primitive recursive models). But since our construction works with maximal consistent sets of formulas and selecting (or constructing) a specific maximal consistent set requires decidability, the arithmetic theory (PA in our case) needs to know that ΣILM is decidable, which we, at this point, did not even show for ourselves.

Therefore we will show ΣILM to be modally complete with respect to a slightly modified semantics in which we have translated some of the frame properties into the forcing relation.

3.3 Modal semantics

3.3.1 A complication

As mentioned, simply taking the intersection of all the ΣL- and ILM-frames will not give the finite model property. Let us show this assertion.

Fact 3.8. *There exists an* IL*-formula which has an infinite* ILM*-model with the additional property $aS_tb \wedge cS_bd \Rightarrow cS_ad$, but does not have a finite such model.*

Proof. Let A be the formula

$$\Diamond(p \rhd q) \wedge ((p \rhd q) \rhd \neg(p \rhd q)) \wedge (\neg(p \rhd q) \rhd (p \rhd q)).$$

Let $M = \langle W, R, S, V\rangle$ where

1. $W = \{x\} + \{y_0, y_1, y_2, y_3 \ldots\} + \{z_1, z_2, z_3, \ldots\}$,
2. $i \geq 0 \Rightarrow xRy_i$, $i \geq 1 \Rightarrow xRz_i$,
3. If i is even and $i < j$ then y_iRz_j,
4. If i is odd then y_iRz_i and if $i+1 < j$ then y_iRz_j,
5. $i \leq j \Rightarrow y_iS_xy_j$,
6. If i is odd and $j < i$, $z_iS_{y_j}z_{i+1}$,
7. $V(p) = \{z_1, z_3, z_5, \ldots\}$ and $V(q) = \{z_2, z_4, z_6, \ldots\}$,

Then M satisfies all the requirements and forces A in x. A part of M is shown in Figure 5. For clarity some arrows which should be there since R and the S_w's are transitive, $wRa \Rightarrow aS_wa$ and $aS_wb \Rightarrow wRa, b$ are not drawn. It remains to

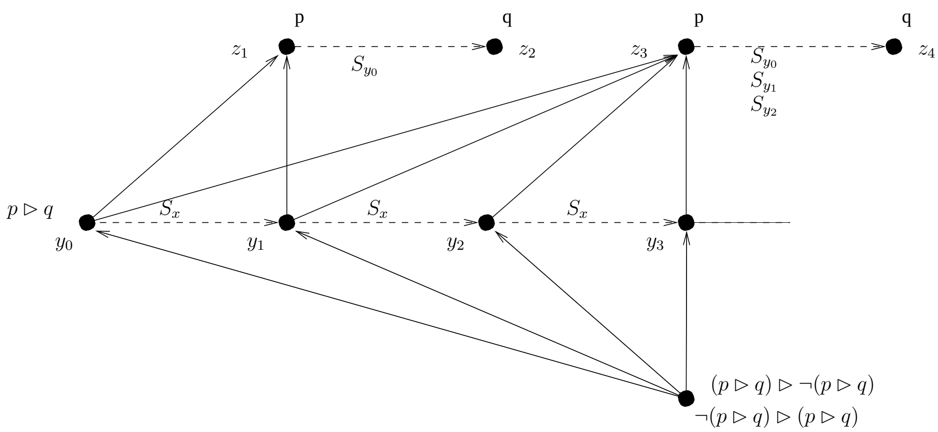


Figure 5: Initial part of M.

show that A does not have a finite model with the required property. So let N be any model satisfying

$$aS_tb \wedge cS_bd \Rightarrow cS_ad. \tag{15}$$

Let $n \in N$ be such that $n \models A$. Let $u = u_0, u_1, u_2, \ldots$ be a sequence of worlds of N such that

a. For all $i \geq 0$, $u_iS_nu_{i+1}$,

b. For all $i \geq 0$ $u_{2i} \models (p \rhd q)$,

c. For all $i \geq 0$ $u_{2i+1} \models \neg(p \rhd q)$.

Such a sequence exists since $n \models A$, it might however be cycling. We claim that this is not so: the set $\{u_i \mid i \geq 0\}$ is infinite. For suppose it is not. Then there exists $i < j$ such that $u_i = u_j$. Assume i is even (the case i is odd goes similarly). We have $u_i S_n u_{i+1} S_n u_j = u_i$. Moreover, since $i+1$ is odd by 3, we can choose some $u_{i+1} R v$ such that $v \models p$ and for no $v S_{u_{i+1}} w$: $w \models q$. By property (M) of ILM-frames $u_i R v$ and thus there exists some $v S_{u_i} w$ such that $w \models q$. But $v S_{u_i} w$ and $u_{i+1} S_n u_i$ imply $v S_{u_{i+1}} w$ by (15), a contradiction. $\dashv$

3.3.2 ΣILM semantics

As we have seen above (Theorem 2.25) taking the S_w relations to be transitive does not harm the ΣL completeness result. It is even easier to see that taking the S_w's reflexive is harmless as well. Modifying the modal completeness proof for ILM as to incorporate $aS_vb \wedge wRv \Rightarrow aS_wb$ (and keeping the model finite) is no problem either but the remaining property: $aS_vb \wedge wS_tv \Rightarrow aS_wb$ is somewhat of a problem as we have shown above. The solution is to transfer this frame property into the forcing relation.

Definition 3.9 (ΣILM-frame). A *ΣILM-frame* is a Veltman frame $\langle W, R, S\rangle$ such that

1. R is transitive and conversely well-founded,
2. for all a, b, c, w:
   (a) $aS_wb \Rightarrow wRa, b$,
   (b) $aS_wbRc \Rightarrow aRc$,
   (c) $wRaRb \Rightarrow aS_wb$,
   (d) $aS_wbS_wc \Rightarrow aS_wc$,
   (e) $aRb \Rightarrow bS_ab$.

Definition 3.10 (ΣILM-model). $\langle W, R, S, V\rangle$ is a *ΣILM-model* if $\langle W, R, S\rangle$ is a ΣILM-frame and V is a mapping PROP $\longrightarrow \mathcal{P}(w)$.

Definition 3.11 (ΣILM forcing relation). For a model M, let $\models_M$ be *the unique relation between W^M and ΣILM-formulas* satisfying:

1. $w \models_M p \Leftrightarrow w \in V^M(p)$,
2. $w \models_M A \rhd B \Leftrightarrow$ for all wRv such that $v \models_M A$ there exists some vS_wu such that $u \models_M B$,
3. $w \models_M \Sigma_1 A \Leftrightarrow$ for all $vS_{w'}u$ such that $w(R \cup \bigcup_{u\in M} S_u)^* w'$: $v \models_M A \Rightarrow u \models_M A$ (see Figure 6),
4. The usual clauses for the boolean connectives.

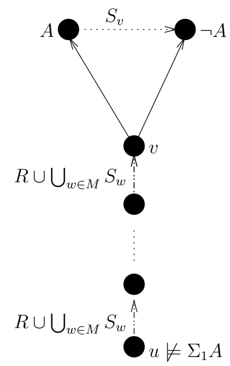


Figure 6: $\Sigma_1 A$ is not forced in u.

As usual we write $M, w \models A$ for $w \models_M A$ or even, when there is no danger of confusion, $w \models A$.

Using this forcing relation we define frame validity as usual: A is valid on a frame F if $w \models_M A$ for any model $M = \langle W^F, R^F, S^F, V\rangle$ and $w \in M$.

3.4 Modal soundness

Theorem 3.12 (Modal Soundness). *If* ΣILM $\vdash A$ *then* $F \models A$ *for any* ΣILM-*frame* F.

Proof. Frame validity is preserved under modus ponens and necessitation. Clearly all propositional tautologies are valid on all frames so it is sufficient to show that all the ΣILM schemata are valid on each frame.

The ΣL part is just like in the modal soundness for ΣL (Theorem 2.7). Where the change in the forcing relation compensates for the change in frame properties.

So it is sufficient to show that the schema M(Σ) and all IL schemata involving the operator $\rhd$ are valid. Below let M be a ΣILM-model and $w, u, v, v' \in M$.

Suppose $w \models \Box(A \to B)$. Suppose wRu and $u \models A$. Then $u \models B$ and thus since $uS_w u$: $w \models A \rhd B$.

Suppose $w \models A \rhd B$ and $w \models B \rhd C$. Suppose wRu and $u \models A$. Then for some v': $uS_w v'$ and $v' \models B$ and thus for some v: $v'S_w v$ and $v \models C$. By transitivity of S_w: $uS_w v$.

Suppose $w \models A \rhd C$ and $w \models B \rhd C$. Suppose wRu and $u \models A \vee B$. Then $u \models A$ or $u \models B$ and in either case there exists some v such that $uS_w v$ and $v \models C$.

Suppose $w \models A \rhd B$ and $w \models \Diamond A$. Then for some u: wRu and $u \models A$. Thus for some v: $uS_w v$ and $v \models B$. Since $S_w \subseteq \{(a, b) \mid wRa, b\}$: $w \models \Diamond B$.

Suppose wRu and $u \models \Diamond A$. Then there exists v such that uRv and $v \models A$ and thus since $wRuRv$ implies $uS_w v$: $w \models \Diamond A \rhd A$.

Suppose $w \models \Sigma_1 C$, $w \models A \rhd B$. Suppose wRu and $u \models A \wedge C$. Then there exists v such that $uS_w v$ and $v \models B$. Since $w \models \Sigma_1 C$ also $v \models C$. $\dashv$

3.5 Modal completeness

3.5.1 Introduction and definitions

In this section we are to prove the modal completeness theorem for the logic ΣILM.

Theorem 3.13 (Modal completeness). *Suppose* ΣILM $\nvdash A$. *Then there exist a* ΣILM*-model* M *and some* $m \in M$ *such that* $m \not\models A$. *Moreover we can take for* m *a root of* M *(that is:* $\forall m' {\in} M\, (m' {\neq} m \rightarrow mRm')$*).*

The method used will be the same as for ΣL above, that is we will be using labeled frames, identify 'problems' in these frames and prove an extension theorem which solves these problems. This method was first applied in [12] to give a modal completeness proof for ILM. Consequently we will be using the same concepts, notations and definitions as before but then translated to the case at hand (e.g. maximal consistent sets are ΣILM-consistent in stead of ΣL-consistent, etc). Also we again use the notion of a relevant set (Definition 2.11).

Applying that method will not go as smoothly as in the case of ΣL. If we simply take the truth definition of $\rhd$ and deny this to create the definition of a problem we will not be able to prove an analogue of Lemma 2.15 (once problems are solved they do not reoccur). The solution is to broaden the notion of a problem in such a way that, although we do a bit too much, we do have that once a problem is solved, it will not return. To this end let us first extend the notion of a (labeled) frame.

Definition 3.14 (Labeled frames). A *labeled frame* is a quintuple $\langle W, R, S, R_e, \nu \rangle$ such that

1. $\langle W, R, S \rangle$ is a frame,
2. ν is a function $W \longrightarrow \mathrm{MCS}$ and
3. $R_e \subseteq \{A \mid A \text{ a } \Sigma\text{ILM-formula}\} \times W \times W$ such that for each B: $R_e^B = \{(u, v) \mid (B, u, v) \in R_e\} \subseteq R$.

Having this extended notion of a frame we define problems as follows:

Definition 3.15 (X-problems). Let X be a set of formulas. An X*-problem* in a labeled frame $\langle W, R, S, R_e, \nu \rangle$ is a world $w \in W$ and a formula $A \in \nu(w) \cap X$ such that one of the following two cases applies:

1. $A = \neg(C \rhd D)$ and there does not exist $wR_e^D v$ with $C \in \nu(v)$.
2. $A = \neg \Sigma_1 B$ and there does not exist $vS_w u$ with $B \in \nu(v)$ and $B \notin \nu(u)$.

Notice that this definition differs from problems in labeled frames in the more natural sense:

$$\neg(B \rhd C) \in \nu^F(w) \text{ but for all } v \text{ for which } wRv \text{ and } B \in \nu^F(v)$$
$$\text{we have a } u \text{ such that } vS_w u \text{ and } C \in \nu^F(u). \quad (16)$$

The $\rhd$ parts in Definition 3.15 and (16) are incomparable. A labeled frame without any problems in the sense of Definition 3.15 might still posses problems in the sense of (16). (And the other way around.) A lot of work will be put in ensuring that Definition 3.15 will be good enough on the frames we will be considering.

Moreover the Σ_1 part in problems as defined above is stronger than we actually need for Theorem 3.13. This stronger version is useful in the arithmetical completeness theorem below so we give it a name.

Definition 3.16 (Strong Σ_1 ΣILM-models). A ΣILM-model is *strong* Σ_1 if whenever $M, w \models \neg\Sigma_1 A$ then there exist u and v with $uS_w v$ such that $u \models A$ and $v \not\models A$.

Theorem 3.17 (Extended modal completeness). *Suppose* ΣILM $\nvdash A$ *then there exist a strong* Σ_1 ΣILM*-model* M *and* $w \in M$ *such that* w *is a root of* M *and* $M, w \not\models A$.

Before we move on let us note that we now do have an analogue of Lemma 2.15.

Lemma 3.18. *Suppose* $F \subseteq G$ *are labeled frames,* $u \in F$*,* A *a formula such that* (u, A) *is not an* X*-problem in* F*. Then* (x, A) *is not an* X*-problem in* G *either.*

Proof. Trivial. $\dashv$

How are we to ensure that our notion of a problem is appropriate? Let us fix a world v. The intention of the relation R_e is to provide for an example for $v \models \neg(B \rhd C)$ whenever $\neg(B \rhd C) \in \nu(v)$.

Let us say that a world w avoids a formula C if for any u for which $wS_v u$ we have $C \notin \nu(u)$. Then we organize things so that, (1) for any D, $vR_e^D w \Rightarrow$ 'w avoids D' and (2) if $\neg(B \rhd C) \in \nu(v)$ then there exists w such that $vR_e^C w$.

We would like to recognize whether a world is avoiding by only considering the world itself. An attempt to this is the notion of critical successor [4].

Definition 3.19 (Critical successor). Let Δ and Γ be MCS's. We say that Δ is a B*-critical successor* of Γ and write $\Gamma \prec_B \Delta$ if $\Gamma \prec \Delta$ and for each A for which $A \rhd B \in \Gamma$: $\neg A, \Box\neg A \in \Delta$.

Lemma 3.20. *Let* Δ*,* Γ *and* Γ' *be* MCS*'s and* B *a formula. If* $\Delta \prec_B \Gamma \prec \Gamma'$ *then* $\Delta \prec_B \Gamma'$.

Proof. $\Delta \prec \Gamma'$ is clear. Suppose $A \rhd B \in \Delta$. Then $\Box\neg A \in \Gamma$ and thus, since $\Gamma \prec \Gamma'$, $\neg A, \Box\neg A \in \Gamma'$. $\dashv$

Part of Definition 3.19 is clear. If $A \in \nu(v)$ and $A \rhd B \in \nu(w)$ then by the truth definition of $\rhd$ (and by anticipating on a truth lemma) there must exist some vS_wu such that $B \in \nu(u)$, precisely what we are trying to avoid. If $\neg\Box\neg A \in \nu(v)$ then for some vRu: $A \in \nu(u)$. By transitivity of R: wRu and thus there exists some u' with uS_wu' and $B \in \nu(u')$. But $wRvRu$ implies vS_wu and so by transitivity of S_w: vS_wu', again not what we want.

So

$$v \text{ avoids } B \Rightarrow \nu(w) \prec_B \nu(v).$$

Can we strengthen this to get a sufficient condition for v to avoid a formula B? Yes:

$$v \text{ avoids } B \Leftrightarrow \forall u (vS_wu \Rightarrow \nu(w) \prec_B \nu(u)).$$

Since $\mathrm{ILM} \vdash B \rhd B$ the sufficiency is clear. Necessity is easily shown by using transitivity of S_w and R.

Let us identify the worlds we have to check whether indeed these R_e^B successors avoid what they should in such a situation.

Definition 3.21 (Critical cones C_x^B). If $\langle W, R, S, R_e, \nu\rangle$ is a labeled frame, B a ΣILM-formula and $x \in W$. Then the B-*critical cone* of x, denoted by C_x^B, is defined to be the smallest set such that:

1. $xR_e^By \Rightarrow y \in C_x^B$,
2. $y \in C_x^B \wedge yS_xz \Rightarrow z \in C_x^B$,
3. $y \in C_x^B \wedge yRz \Rightarrow z \in C_x^B$.

Our main concern will be to ensure that the C_x^B-Critical cones (indeed) lie B-critically above x.

As in the ΣL case there is more than just solving problems (but see the remarks before Definition 3.23). There is the issue of reasonability. In this notion we have incooperated the above considerations on the critical cones. And an additional technical requirement.

Definition 3.22 (Reasonable). A labeled frame F is called *reasonable* if

1. For each B, $x \in F$ and $y \in C_x^B$: $\nu(y)$ lies B-critically above $\nu(x)$,
2. For each x and $B \neq B'$: $C_x^B \cap C_x^{B'} = \emptyset$,
3. $xRy \Rightarrow \nu^F(x) \prec \nu^F(y)$,
4. $xS_zy \Rightarrow \nu^F(x) \subseteq_{\Sigma_1, \nu^F(z)} \nu^F(y)$.

If we compare the notions introduced in this section to the corresponding notions in the ΣL case we see a difference.

Suppose we want to show that a labeled, reasonable, ΣL-frame F without problems satisfies a (ΣL) truth lemma: '$v \models A \Leftrightarrow A \in \nu^F(v)$'. If we proceed by induction on A then the nonexistence of problems handles the $\neg\Box$ and $\neg\Sigma_1$ cases and the reasonability of F handles the $\Box$ and Σ_1 cases. The possibility of such a treatment lies in the quantifier complexity of the $\neg\Box$ and $\neg\Sigma_1$ truth definitions ($\exists$) and the quantifier complexity of the $\Box$ and Σ_1 truth definitions ($\forall$).

The formulation of truth of an IL-formula has, however, quantifier complexity $\forall\exists$ and its negation, thus, $\exists\forall$. Via the introduction of the relation R_e we will manage to handle the $\neg(A \rhd B)$ truth definition as if it has quantifier complexity $\exists$. Luckily a property of a labeling function ν with quantifier complexity $\forall\exists$ behaves quite well under frame extensions (Lemma 3.25 and Lemma 3.24). So problems in frames involving a formula $B \rhd C$, called defects, are handled 'directly'.

Definition 3.23 (X-defects). An X-*defect* in a labeled frame $\langle W, R, S, R_e, \nu\rangle$ is a world $w \in W$ together with a formula $C \rhd D \in \nu(w) \cap X$ such that for some wRv with $C \in \nu(v)$ there does not exists a $vS_w u$ with $D \in \nu(u)$. In this case we say that (w, A) is an X-*defect* (or simply *defect* when X is understood) w.r.t. v.

Lemma 3.24. *If $F \subseteq G$ are labeled frames, B a formula and $x, y \in F$ such that (x, B) is a defect w.r.t. y in G then (x, B) is a defect w.r.t. y in F.*

Proof. Trivial. $\dashv$

Lemma 3.25. *Let $F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots$ be a chain of labeled frames. If for each $i \geq 0$ F_i is reasonable then so is $F = \bigcup_{i\geq 0} F_i$.*

Proof. Trivial. $\dashv$

And, of course, besides ensuring a truth lemma we need to make sure that our model satisfies the ΣILM-frame properties:

Definition 3.26 (ΣILM-closure). If $F = \langle W, R, S\rangle$ is a finite frame such that R is conversely well-founded. Then the ΣILM-*closure* of F is defined as the intersection of all ΣILM-frames with domain W (the domain of F) which extend F.

Lemma 3.27. *Let $F = \langle W, R, S\rangle$ be a finite frame such that R is conversely well founded. If G is the* ΣILM *closure of F then:*

1. *G is a* ΣILM*-frame,*
2. *If $(x, y) \in R^G$ or, for some t, $(x, y) \in S_t^G$. Then there exists a z such that $(z, y) \in R^F$ and $(x, z) \in (R^F \cup \bigcup_{x\in F} S_x^F)^*$,*
3. *G is reasonable if F is.*

Proof. Since the intersection of any set of ΣILM-frames is a ΣILM-frame and there exists at least one frame extending F which satisfies all the ΣILM-frame properties the first assertion is trivial. Knowing this it is easy to see that the ΣILM closure is equal to the union of a chain $F_0 \subseteq F_1 \subseteq \cdots$ which satisfies[10]:

[10] Since F is finite any such chain will do. If we allow F to be infinite we need to be more careful.

1. $F_0 = F$,
2. For $i \geq 1$:
   (a) $xR^{F_i}yR^{F_i}z$ but not $xR^{F_i}z$ and $F_{i+1} = \langle W, R^{F_i} + \{(x,z)\}, S^{F_i}\rangle$ or
   (b) $xS_t^{F_i}yRz$ but not $xR^{F_i}z$ and $F_{i+1} = \langle W, R^{F_i} + \{(x,z)\}, S^{F_i}\rangle$ or
   (c) $xS_t^{F_i}yS_t^{F_i}z$ but not $xS_t^{F_i}z$ and $F_{i+1} = \langle W, R^{F_i}, S^{F_i} + \{(x,t,z)\}\rangle$ or
   (d) $xR^{F_i}y$ but not $yS_x^{F_i}y$ and $F_{i+1} = \langle W, R^{F_i}, S^{F_i} + \{(x,y,y)\}\rangle$.
   (e) $xR^{F_i}yR^{F_i}z$ but not $yS_x^{F_i}z$ and $F_{i+1} = \langle W, R^{F_i}, S^{F_i} + \{(y,x,z)\}\rangle$ or
   (f) F_i is a ΣILM-frame and $F_{i+1} = F_i$.

Now one easily shows that 2 holds for each F_i in such a chain and therefore for G as well.

To show 3 only that the ΣILM-closure satisfies 3.22-1 needs a proof. We show that 3 holds for each F_i with induction on i. The critical cones in F_0 and F_1 are equal to those in F so assume $i \geq 1$. Suppose we are in situation 2e. Since successors of critical successors are themselves critical successors (Lemma 3.20) F_{i+1} satisfies 3.22-1 if F_i does. The other cases are handled similarly. ⊣

3.5.2 Tools

As in the ΣL case we need some mathematical facts on maximal consistent sets. In one of the proofs below we will be using a well-known equivalent (equivalent in ZF set theory) of the axiom of choice known as 'Zorn's lemma', which reads as follows: If $(X, \leq)$ is a partial order such that any chain has an upper bound then $(X, \leq)$ has a maximal element, that is: there exists some $x \in X$ such that for all $y \in X$: $x \leq y \Rightarrow x = y$. See for example [14].

Lemma 3.28. *Suppose Δ is an* MCS *and $\neg\Sigma_1 A \in \Delta$. Then there exist* MCS*'s Γ_0 and Γ_1 such that $\Gamma_0 \subseteq_{\Sigma_1,\Delta} \Gamma_1$, $A \in \Gamma_0$, $A \notin \Gamma_1$ and $\Sigma_1 A \in \Gamma_0 \cap \Gamma_1$.*

Proof. This lemma has been proven in the context of ΣL above (Lemma 2.20). Exactly the same proof works here. ⊣

Lemma 3.29. *If Δ is an* MCS *and $\neg(B \rhd C) \in \Delta$ then there exists an* MCS *Γ such that $\Delta \prec_C \Gamma$, $B \in \Gamma$ and $\Box\neg B \in \Gamma$.*

Proof. Let Δ, B, C be as in the condition of the lemma. Then the set

$$\{D, \Box D \mid \Box D \in \Delta\} + \{\Box\neg A, \neg A \mid A \rhd C \in \Delta\} + \{B, \Box\neg B\} \tag{17}$$

is consistent. For suppose it is not. Then there exist $D_0, \ldots, D_k, A_0, \ldots, A_l$ such that:

$$\vdash \bigwedge_{0\leq i\leq k} (D_i \wedge \Box D_i) \wedge \bigwedge_{0\leq i\leq l} (\neg A_i \wedge \Box\neg A_i) \wedge B \wedge \Box\neg B \rightarrow \bot.$$

By some standard GL reasoning:

$$\vdash (\bigwedge_{0\leq i\leq k} \Box D_i) \rightarrow \Box(B \wedge \Box\neg B \rightarrow \bigvee_{0\leq i\leq l} (A_i \vee \Diamond A_i)),$$

and therefore by IL Axiom 3:

$$\vdash (\bigwedge_{0\leq i\leq k} \Box D_i) \rightarrow (B \wedge \Box\neg B \rhd \bigvee_{0\leq i\leq l} (A_i \vee \Diamond A_i)),$$

and thus applying Lemma 3.1 2-3 and IL Axiom 3 we conclude

$$\vdash (\bigwedge_{0\leq i\leq k} \Box D_i) \rightarrow B \rhd \bigvee_{0\leq i\leq l} A_i.$$

For each $0 \leq i \leq k$: $\Box D_i \in \Delta$ and thus $B \rhd \bigvee_{0\leq i\leq l} A_i \in \Delta$. Using that for each $0 \leq i \leq l$: $A_i \rhd C \in \Delta$ we conclude $B \rhd C \in \Delta$, a contradiction.

So (17) is consistent and we can take for Γ any MCS extending it. ⊣

Lemma 3.30. *If Δ is an* MCS*, $B \rhd C \in \Delta$ and there exists Γ such that $B \in \Gamma$ and for some D: $\Delta \prec_D \Gamma$ then there exists an* MCS *Γ' such that $\Delta \prec_D \Gamma'$, $\Gamma \subseteq_{\Sigma_1,\Delta} \Gamma'$ and $C \in \Gamma'$. Moreover among such Γ' there exists a $\subseteq_{\Sigma_1,\Delta}$-maximal one.*

Proof. We first show the existence of such a Γ' and then the existence of a $\subseteq_{\Sigma_1,\Delta}$-maximal one.

Suppose Δ, B, C, Γ and D satisfy the conditions of the lemma. We claim that

$$\{E, \Box E \mid \Box E \in \Delta\} + \{\Box\neg A, \neg A \mid A \rhd D \in \Delta\} + \{C\} \tag{18}$$

is consistent. For if this is not the case we can reasoning as in Lemma 3.29 and conclude $B \rhd D \in \Delta$. But since $\Delta \prec_D \Gamma$ this then implies $\neg B \in \Gamma$, a contradiction. Next we claim that

$$(18) + \{D \in \Gamma \mid \Sigma_1 D \in \Delta\} \tag{19}$$

is consistent. For suppose it is not. Then for some $D_0, \ldots, D_k \in \{D \in \Gamma \mid \Sigma_1 D \in \Delta\}$ the set

$$(18) + \{D_0, \ldots, D_k\} \tag{20}$$

is inconsistent. But $\vdash \Sigma_1 D_0 \wedge \cdots \wedge \Sigma_1 D_k \rightarrow \Sigma_1(D_0 \wedge \cdots \wedge D_k)$ so $\Sigma_1(D_0 \wedge \cdots \wedge D_k) \in \Delta$. Therefore replacing B by $B \wedge D_0 \wedge \cdots \wedge D_k$ and C by $C \wedge D_0 \wedge \cdots \wedge D_k$ in the above argument which shows that (18) is consistent then shows that (20) is consistent, a contradiction and thus (19) is consistent. Now for Γ' we can take any MCS which extends (19).

Now, for the existence of a $\subseteq_{\Sigma_1,\Delta}$-maximal MCS among those Γ''s, let

$$\Gamma'_0 \subseteq_{\Sigma_1,\Delta} \Gamma'_1 \subseteq_{\Sigma_1,\Delta} \Gamma'_2 \subseteq_{\Sigma_1,\Delta} \cdots \tag{21}$$

be a chain of MCS's extending (19). Let $\Gamma'_\infty = \bigcup_{0\le i}\{E \in \Gamma'_i \mid \Sigma_1 E \in \Delta\}$. Then

$$\Gamma'_\infty + (19) \tag{22}$$

is consistent. For if this is not the case then for some finite subset Γ'' of Γ'_∞ the set $\Gamma'' + (19)$ is inconsistent. But this set is a subset of some Γ'_i, a contradiction. Now any MCS extending (22) will be an MCS extending (19) and a $\subseteq_{\Sigma_1,\Delta}$-upper-bound for the chain (21). Therefore by Zorn's lemma there exists a Γ'_{max} extending (19) such that for all Γ' extending (19) we have $\Gamma'_{max} \subseteq_{\Sigma_1,\Delta} \Gamma' \Rightarrow \Gamma' = \Gamma'_{max}$. $\dashv$

The $\subseteq_{\Sigma_1,\Delta}$-maximality of Γ' in the above lemma is used to show the existense of a finite countermodel for an unprovable formula. In [12] $\subset_\Box$-maximal sets are used to do the same for ILM. ($\Delta \subset_\Box \Gamma \Leftrightarrow \{\Box C \mid \Box C \in \Delta\} \subseteq \Gamma\}$.)

3.5.3 Extension theorem

Definition 3.31. if x is a node in a labeled frame then put

$$|x|_A = \#\{\neg\Sigma A \in \nu(x) \cap \mathcal{R}_A\} + \#\{\neg\Box\neg A \in \nu(x) \cap \mathcal{R}_A\}$$

With the above definition we can bound the number of problems in a world in the sense that if $|x|_A = 0$ then x does not have any $\mathcal{R}_A$-problems. Moreover if the frame in question is reasonable:

$$xRy \Rightarrow |x|_A \ge |y|_A. \tag{23}$$

Using $\#\{\neg(A \rhd B) \in \nu(x) \cap \mathcal{R}_A\}$ instead of $\#\{\neg\Box\neg A \in \nu(x) \cap \mathcal{R}_A\}$ might seem more logical but then, since $\Sigma\mathrm{ILM} \nvdash A \rhd B \to \Box(A \rhd B)$, we cannot guarantee (23).

Theorem 3.32 (Extension Theorem). *Let X be a set of ΣILM-formulas. If F is a reasonable ΣILM-frame without any X-defects and with some problem (x, A). Then there exists a reasonable ΣILM-frame G without any X-defects such that $F \subseteq G$ and G does not have the problem (x, A).*

Moreover if F and X are finite then so is G. Also if B is a formula such that $A \in \mathcal{R}_B$ then for each $v \in G - F$: $|v|_B < |x|_B$.

Proof. Let F be as in the conditions of the theorem.

First, let $\mathcal{Q}$ be some set such that $W^F \subseteq \mathcal{Q}$ and $\mathcal{Q} - W^F$ is countable infinite, and $<$ be some well-ordering on $\mathcal{Q} \times \{A \mid A \text{ a } \Sigma\text{ILM-formula}\} \times \mathcal{Q}$.

Now let (x, A) be some problem. G will be the union of a chain $G_0 \subseteq G_1 \subseteq \cdots$, of reasonable labeled frames defined inductively as follows:

- For the definition of G_0 we distinguish the following cases:
  - If $A = \neg\Sigma_1 B$ then

    $$G_0 = \langle W^F + \{z_0, z_1\}, R^F + \{(x, z_0), (x, z_1)\}, S^F + \{(x, z_0, z_1)\}, R^F_e, \nu^F + \{(z_0, \Delta_0), (z_1, \Delta_1)\}\rangle$$

    for some $z0, z1 \in \mathcal{Q} - W^F$ and MCS's Δ_0 and Δ_1 such that $B \in \Delta_0$, $B \notin \Delta_1$, $\Sigma_1 B \in \Delta_0 \cap \Delta_1$ and $\Delta_0 \subseteq_{\Sigma_1,\nu^F(x)} \Delta_1$.

- Else if $A = \neg(B \rhd C)$ then

  $$G_0 = \langle W^F + \{z\}, R^F + \{(x,z)\}, S^F, R_e^F + \{(C,x,z)\}, \nu^F + \{(z,\Delta)\}\rangle$$

  for some $z \in \mathcal{Q} - W^F$ and an MCS Δ such that Δ is a C-critical successor of $\nu^F(x)$, $\Box\neg B \in \Delta$ and $B \in \Delta$.

- If G_j has been defined. Then

  1. If G_j is not a ΣILM-frame then let G_{j+1} be the ΣILM-closure of G_j,
  2. else if no defects in G_j exist put $G_{j+1} = G_j$,
  3. else let $(z, C \rhd D, y')$ be some $<$-minimal element such that $(z, C \rhd D)$ is a defect in G_j w.r.t. y'.
     Let us call an MCS Δ good if it satisfies:
     - $\nu^{G_j}(y') \subseteq_{\Sigma_1,\nu(z)} \Delta$,
     - $\nu(z) \prec \Delta$,
     - $D \in \Delta$,
     - if for some B, $y' \in C_z^B$. Then Δ lies B-critically above $\nu^{G_j}(z)$.

     Now distinguish the following two cases:
     (a) If there exists some $y \in G_j - G_0$ such that
        i. $\nu^{G_j}(y)$ is good,
        ii. zRy,
        iii. if for some B: $y' \in C_z^B$ then $y \in C_z^B$.

        then put $G_{j+1} = \langle W^{G_j}, R^{G_j}, S^{G_j} + \{(z,y',y)\}, R_e^{G_j}, \nu^{G_j}\rangle$.
     (b) Else let $y'' \in \mathcal{Q} - G_j$ and Δ some MCS such that
        i. Δ is good,
        ii. Δ is maximal w.r.t. $\subseteq_{\Sigma_1,\nu^{G_j}(z)}$ among the MCS's which are good,

        and put $G_{j+1} = \langle W^{G_j} + \{y''\}, R^{G_j} + \{(z,y'')\}, S^{G_j} + \{(z,y',y'')\}, R_e^{G_j}, \nu^{G_j} + (y'',\Delta)\rangle$.

In what follows, to enhance readability, I will drop all superscripts on the labeling functions ν. So instead of ν^{G_j} or ν^{G_i} I will just write ν. Since if $j > i$ then G_j is an extension of G_i there never is any danger of misunderstanding. Alternatively one could think every ν superscripted with G.

Lemma. *For each $j \geq 0$: G_j is well-defined and reasonable.*

Proof. We show simultaneously that each G_j is well-defined and reasonable by induction on j.
G_0 is clearly well-defined. G_0 is reasonable since F is.
Now suppose G_j is reasonable (and well-defined). We have to distinguish three cases:

1. G_{j+1} is the ΣILM-closure of G_j,

2. G_{j+1} is defined out of G_j by applying case 3a or

3. by applying case 3b.

If we are in Case 1 then G_{j+1} is clearly well defined and G_{j+1} is reasonable by Lemma 3.27.
Case 2 is trivial.

So suppose we are in Case 3. Assume that for some B, $y' \in C_z^B$. Such a B is unique, so we can find Δ by Lemma 3.30. And thus G_{j+1} is well defined. The newly added y'' is in C_z^B and in no other critical cone of z. And since $\nu(y'')$ lies B-critically above $\nu(z)$ this implies that G_{j+1} is reasonable. The case that for no B, $y' \in C_z^B$ goes similarly. $\dashv$

The rest of the proof of Theorem 3.32 is given by the following three claims.

Claim 1. G is reasonable, a ΣILM-frame, does not have any X-defects and does not have the problem (x, A).

Proof. G does not have any X-defects since if $x, y \in G$ then for some j: $x, y \in G_j$, but then if (x, B) is a defect w.r.t. y in G then by Lemma 3.24 (x, B) is a defect w.r.t. y in G_j. If $n = \#\{z \in \mathcal{Q} \times \{A \mid A \text{ a } \Sigma\text{ILM-formula}\} \times \mathcal{Q} \mid z < (x, B, y)\}$ then for some $k \leq n+1$: (x, B) is not a defect w.r.t. y in G_{j+k} and thus, again by Lemma 3.24 it is not a defect in G either. A contradiction.

Furthermore by Lemma 3.25 G is reasonable since each G_j is. And finally by Lemma 3.18 G does not have the problem (x, A) since G_0 does not have this problem and $G_0 \subseteq G$. $\dashv$

Claim 2. if D is such that $A \in \mathcal{R}_D$ then for each $v \in G - F$, $|v|_D < |x|_D$.

Proof. We show by induction on j that if $v \in G_j - F$ then $|v|_D < |x|_D$. If $j = 0$ then xRv and thus $|v|_B \leq |x|_B$.

If $A = \neg\Sigma_1 A'$ then $\Sigma_1 A' \in \nu(v)$. If $A = \neg(B \rhd C)$ then $\Box\neg B \in \nu(v)$ since ILM $\vdash B \rhd B$ and $\nu(v)$ is chosen to be a B-critical successor of $\nu(x)$. In either case if $A \in \mathcal{R}_D$ we conclude: $|v|_D < |x|_D$.

The inductive step is handled trivially using $a < b \leq c \Rightarrow a < c$ and the fact (for reasonable frames): $uS_w v \Rightarrow |u|_D \geq |v|_D$. $\dashv$

Claim 3. if F and X are finite then so is G.

Proof. First a word on terminology. Using the same (variable and formula) names as in the definition of the chain. If G_{j+1} is defined out of G_j by applying Case 3a then we will say that: 'The defect $(z, C \rhd D)$ w.r.t y' in G_j is solved by $y'S_z y$ by applying Case 3a.' Similarly if G_{j+1} is defined out of G_j by applying Case 3b we say that: 'The defect $(z, C \rhd D)$ w.r.t y' in G_j is solved by $y'S_z y''$ by applying Case 3b.' Additionally when, for instance, $z, C \rhd D, y'$ and y'' are understood or nonimportant we may just say: 'The defect in G_j is solved by applying Case 3b.'

Let us assume, for a contradiction, that F and X are finite and that G is infinite. Then there would exist a sequence:

$$y_0 S_{z_0} y_1 S_{z_1} y_2 \dots$$

and a sequence of formulas:

$$C_0 \rhd D_0, C_1 \rhd D_1, \dots$$

and a sequence $j_0 < j_1 < \dots$ such that the defect $(z_i, C_i \rhd D_i)$ w.r.t. y_i in G_{j_i} is solved by $y_i S_{z_i} y_{i+1}$ by applying Case 3b. See Figure 7.

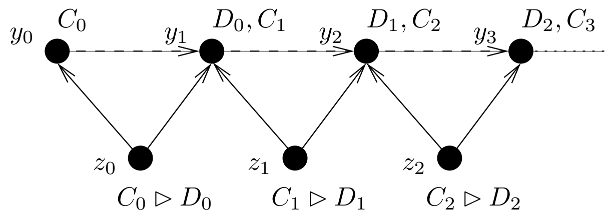


Figure 7: An infinite chain of solved defects.

Claim 3a. There exist $k < l$, $z, C \rhd D$ such that

- $z = z_l = z_k$,
- $C \rhd D = C_l \rhd C_l = C_k \rhd D_k$,
- for any B: $y_{l+1} \in C_z^B$ iff $y_{k+1} \in C_z^B$.

Proof. For each $i \geq 0$: $y_i \in G - F$ since, as one easily checks, defects only occur w.r.t. worlds in $G - F$. Since for worlds $v \in G - F$ pairs (u, v) are in $R^G - R^F$ only for worlds $u \in F$ this implies that for all i: $z_i \in F$.

Now each of those y_i can be colored by three of its properties:

1. The world $z_i \in F$,
2. the formula $C_i \rhd D_i \in \nu(z_i) \cap X$,
3. a formula B such that $y_i \in C_{z_i}^B$ or $\perp$ if no such B exists.

Since F is finite there are only finitely many possibilities for Property 1. There are also finitely many possibilities for Property 3 since for each z_i and $B \in$ MCS $C_{z_i}^B$ is nonempty iff there exists some y such that $z_i R_e^B y$. Clearly R_e^G is finite so only finitely many C_z^B are nonempty. X is finite so there can be only finitely many possibilities for Property 2 as well. Conclusion: there are only finitely many colors and there exists $k < l$ such that y_k and y_l have the same color $(z, C \rhd D, B)$. $\dashv$

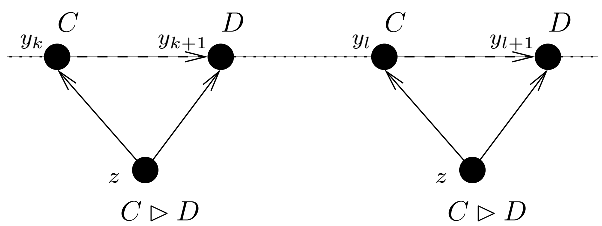


Figure 8: The defect $(z, C \rhd D)$ occurs twice.

Fix $k < l$, z and $C \rhd D$ as given by Lemma 3a. See figure 8.

We are going to show that we can 'link back' y_l to y_{k+1}. The only non trivial thing we need for this is $\nu(y_l) \subseteq_{\Sigma_1,\nu(z)} \nu(y_{k+1})$. We are going to show $\nu(y_{k+1}) = \nu(y_{l+1})$. Which is clearly sufficient.

Claim 3b. $\nu(y_{k+1}) \subseteq_{\Sigma_1,\nu(z)} \nu(y_{l+1})$.

Let us first see why this claim finishes the proof of Claim 3. By 3(b)ii $\nu(y_{k+1})$ is chosen to be maximal w.r.t. $\subseteq_{\Sigma_1,\nu(z)}$ among the MCS's Δ with the properties:

1. $D \in \Delta$,
2. $\nu(y_k) \subseteq_{\Sigma_1,\nu(z)} \Delta$ and
3. if for some B: $y_k \in C_z^B$ then Δ lies B-critically above $\nu(z)$.

Clearly $\nu(y_{l+1})$ satisfies properties 1 and 3. Claim 3b implies that $\nu(y_{l+1})$ satisfies 2.

So $\nu(y_{k+1})$ is a $\subseteq_{\Sigma_1,\nu(z)}$-maximal element of a class $\nu(y_{l+1})$ belongs to and thus by Claim 3b: $\nu(y_{l+1}) = \nu(y_{k+1})$. This together with the fact $zR^{G_{j_l}}y_{k+1}$ implies that the defect $(z, C \rhd D, y_l)$ in G_{j_l} is solved by applying Case 3a. A contradiction.

So indeed Claim 3b suffices.

Proof of Claim 3b. Let $\overline{S}^F = \bigcup_{w \in F} S_w^F$. We show by induction on $l - h$:

$$\text{If } k \leq h \leq l \text{ then } z_l(R^F \cup \overline{S}^F)^* z_h. \tag{24}$$

See Figure 9.

If $h = l$ the statement is trivial. So suppose the statement holds for $h + 1$. It is sufficient to show

$$z_{h+1}(R^F \cup \overline{S}^F)^* z_h. \tag{25}$$

Lemma. *for all $i \geq 0$ and $u \in F$:*

$$uR^{G_i}y_{h+1} \Rightarrow u(R^F \cup \overline{S}^F)^* z_h.$$

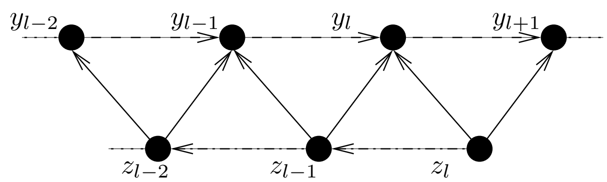


Figure 9: $z_{i+1}(R^F \cup \bigcup_{w\in F} S^F_w)^* z_i$.

Proof. Induction on i. If $i \leq j_h+1$ we are done since $uR^{G_{j_h+1}}y_{h+1} \Leftrightarrow u = z_h$. So assume the statement for some $i \geq j_h+1$ and suppose $uR^{G_{i+1}}y_{h+1}$. If $uR^{G_i}y_{h+1}$ then we are done so suppose this is not the case. Then since $y_{h+1} \in G_i$ it must be the case that G_{i+1} is the ΣILM-closure of G_i, since this is the only way that 'R relations are added between existing worlds'. By Lemma 3.27 Item 2 there exists some $v \in G_i$ such that $vR^{G_i}y_{h+1}$ and $u(R^F \cup \overline{S}^F)^*v$. By (IH) $v(R^F \cup \overline{S}^F)^* z_h$. Conclusion: $u(R^F \cup \overline{S}^F)^* z_h$. ⊣

By Lemma 3.5.3 and since $z_{h+1}R^{G_{j_{h+1}}}y_{h+1}$ we conclude (25) and thus also (24).

Now suppose $\Sigma_1 D \in \nu(z)$, $D \in \nu(y_{k+1})$. Since $z = z_l$: $\Sigma_1 D \in \nu(z_l)$. Formulas of the form $\Sigma_1 D$ are preserved along S_u (by reasonability) and R (by reasonability and since $\Sigma\mathrm{ILM} \vdash \Sigma_1 A \to \Box\Sigma_1 A$). And thus by (24)

$$\text{for each } k < h \leq l \quad \Sigma_1 D \in \nu(z_h).$$

So since

$$\nu(y_{k+1}) \subseteq_{\Sigma_1,\nu(z_{k+1})} \nu(y_{k+2}) \subseteq_{\Sigma_1,\nu(z_{k+2})} \ldots \subseteq_{\Sigma_1,\nu(z_l)} \nu(y_{l+1}),$$

it follows that

$$\text{for each } k \leq h \leq l \quad D \in \nu(y_{h+1}).$$

end of proof of Claim 3b. ⊣

end of proof of Claim 3. ⊣

end of proof of Theorem 3.32. ⊣

Proof of Theorem 3.13. Let A be a formula as in the assumption of the theorem: $\Sigma\mathrm{ILM} \nvdash A$. We define a labeled ΣILM-frame F as the union of a chain $F_0 \subseteq F_1 \subseteq \cdots$ of finite, reasonable ΣILM-frames, all of which are without any $\mathcal{R}_A$-defects, inductively as follows:

- $F_0 = \langle \{w\}, \emptyset, \emptyset, \emptyset, \{(w, \Delta_A)\}\rangle$, where Δ_A is some MCS containing $\neg A$.
- Suppose F_j has been defined.
  - If F_j does not have any $\mathcal{R}_A$-problems then put $F_{j+1} = F_j$,

– else let (x, A) be some $\mathcal{R}_A$-problem in F_j and let F_{j+1} be a frame as given by the extension theorem (Theorem 3.32) (with $\mathcal{R}_A$ for X and (x, A) for (w, A)).

As F is the union of reasonable frames F itself is reasonable (Lemma 3.25). Similarly as the union of a chain of frames without any $\mathcal{R}_A$-defects, F itself does not have any $\mathcal{R}_A$-defects.

One shows as in the ΣL case that F is finite, which implies that F is a ΣILM-frame and does not contain any $\mathcal{R}_A$-problems.

Let M be the model defined out of F by putting $V(p) = \{v \mid p \in \nu^F(v)\}$. As in the ΣL case: $M, w \models \neg A$. ⊣

3.6 Failure of interpolation

In this section we extend 'failure of interpolation for ΣL' (Section 2.4) to ΣILM. The definition of interpolant (Definition 2.24) clearly translates to ΣILM so we simply state the theorem:

Theorem 3.33. ΣILM *does not have interpolation.*

First let us proof two preliminary lemmas.

Lemma 3.34. *For all* ΣILM*-formulas* $A, B, D = D(p)$*:*

$$\Sigma\mathrm{ILM} \vdash \boxdot(A \leftrightarrow B) \to (D(A) \leftrightarrow D(B)).$$

Proof. By the modal completeness theorem it is sufficient to show that $\boxdot(A \leftrightarrow B) \to (D(A) \leftrightarrow D(B))$ is forced in any root of a ΣILM-model.

So let M be a ΣILM-model and let m be a root. Suppose $m \models \boxdot(A \leftrightarrow B)$. Since m is a root of M: $A \leftrightarrow B$ is forced in any world of the model so trivially $D(A)$ is forced in m iff $D(B)$ is forced in m. ⊣

Lemma 3.35. *Let* $D = D(p)$ *be a* ΣILM*-formula such that* p *only occurs under the scope of a* $\rhd$ *or* Σ_1*. Then for any two formulas* A *and* B*:*

$$\Sigma\mathrm{ILM} \vdash \Box(A \leftrightarrow B) \to (D(A) \leftrightarrow D(B))$$

Proof. Induction on D. If $D = q$ then $p \neq q$ and thus $D(A) = D(B)$. Truth functional cases are trivial.

Assume $D = D_0(p) \rhd D_1(p)$. Again we use the modal completeness theorem. Let w be a world in a ΣILM-model for which we have $w \models \Box(A \leftrightarrow B)$. Then for each v for which wRv we have $v \models \boxdot(A \leftrightarrow B)$ and thus by Lemma 3.34 above for each v for which wRv:

$$v \models D_0(A) \leftrightarrow D_0(B),$$
$$v \models D_1(A) \leftrightarrow D_1(B).$$

Now trivially by the definition of $\rhd$: $w \models D_0(A) \rhd D_1(A) \leftrightarrow D_0(B) \rhd D_1(B)$.

The case $D = \Sigma_1 D_0(p)$ goes similarly. ⊣

Corollary 3.36. *If $D = D(p_0, \dots, p_k, q_0, \dots, q_r)$ is a* ΣILM*-formula such that each p_i and q_j only occurs under the scope of a $\rhd$ or Σ_1 then:*

$$\Sigma\mathrm{ILM} \vdash \Box\Box\bot \to \{D(A_0 \rhd B_0, \dots, A_k \rhd B_k, \Sigma_1 C_0, \dots, \Sigma_1 C_r) \leftrightarrow D(\top, \dots, \top, \top, \dots, \top)\}.$$

Proof. $\Sigma\mathrm{ILM} \vdash \Box\bot \to ((p \rhd q) \wedge \Sigma_1 r)$ so

$$\Sigma\mathrm{ILM} \vdash \Box\Box\bot \to \Box(p \rhd q \leftrightarrow \top) \text{ and } \Sigma\mathrm{ILM} \vdash \Box\Box\bot \to \Box(\Sigma_1 r \leftrightarrow \top).$$

So the corollary follows by repeated application of Lemma 3.35. ⊣

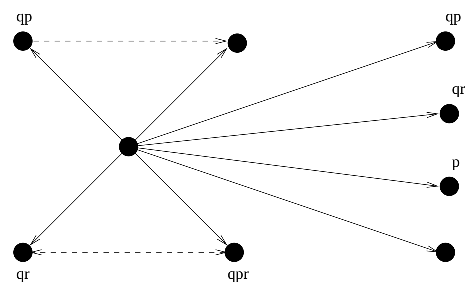


Figure 10: M

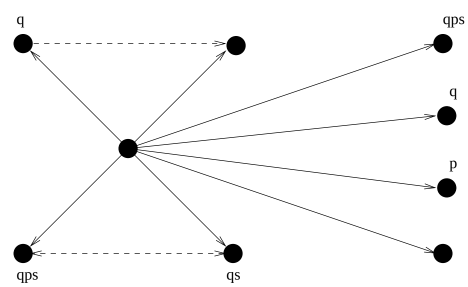


Figure 11: M'

Proof of Theorem 3.33. Just as in the proof of Theorem 2.25 let

$$A(p, q, s) = A_s =_{\text{def}} \neg\Sigma_1 q \wedge \Sigma_1 s \wedge \Box(s \to q) \wedge \Box(p \wedge q \to s),$$
$$B(p, q, r) = B_r =_{\text{def}} \neg\Sigma_1 q \wedge \Sigma_1 r \wedge \Box(r \to q) \wedge \Box(\neg p \wedge q \to r).$$

As was shown in the proof of Theorem 2.25 $\{A_s, B_r\}$ is inconsistent.

To show that no interpolant exists consider the Figures 11 and 10. Figure 10 shows a model M and Figure 11 a model M'. If we say that a formula is forced in M, resp. M', we mean that it is forced in the center point of M, resp M'. M' is a model of A_s and M is a model of B_r.

In a very similar way as in the proof of Theorem 2.25 one shows that these two models force the same ΣL-formulas. We will first show that they force the same IL-formulas as well and then combine these two facts to show that they force the same ΣILM-formulas.

To show that they force the same IL-formulas it is sufficient to give an IL-bisimulation (see Definition 2.30). It is easy to check that Z as given by:

Definition 3.37. Let Z be the binary relation between M and M' such that $(m, m') \in Z$ iff one of the following cases applies

- m and m' are both center points.
- m and m' both are not center points and:

$$m \in V(p) \Leftrightarrow m' \in V'(p) \text{ and}$$
$$m \in V(q) \Leftrightarrow m' \in V'(q).$$

is an IL-bisimulation w.r.t. $\{p, q\}$.

Now in order to show that these two models force the same ΣILM formulas let D be some ΣILM-formula which is forced in one of them. Note that both models are ΣILM-models and force $\Box\Box\bot$ and thus by Corollary 3.36 we can assume w.l.o.g. that the operators $\rhd$ and Σ_1 do not occur nested: D is a boolean combination of ΣL and IL-formulas. But since both models force the same of such formulas they either both force D or both do not. $\dashv$

4 Arithmetical interpretation

In this section we are going to give arithmetical meaning to the logic ΣILM.

Now capital letters like A, B and C can denote ΣILM-formulas as well as first order formulas, with identity, in the language of PA: $\langle +, \times, 0, 1 \rangle$.

We assume a coding of the syntax of PA in PA [2]. If A is a formula we denote by $\ulcorner A \urcorner$ its code. With $\dot{\neg}$ we denote the primitive recursive function such that for each formula A: $\dot{\neg}\ulcorner A \urcorner = \ulcorner \neg A \urcorner$. Similar conventions hold for the other boolean connectives. Moreover we assume a formalization of provability in PA: a predicate $\Box(x)$ such that for any sentence A:

$$\mathrm{PA} \vdash A \Leftrightarrow \mathrm{PA} \vdash \Box(\ulcorner A \urcorner). \tag{26}$$

Bold face characters like $\mathbf{n}$ and $\mathbf{w}$ denote fixed (standard) natural numbers. We do not make a distinction between natural numbers and numerals. Normal characters like n and w are (just) variables.

Using a primitive recursive function $\lambda xyz.\mathrm{sub}(x, y, z)$ such that for each formula $A(x)$ and $\mathbf{n} \in \omega$: $\mathrm{sub}(\ulcorner A(x) \urcorner, \ulcorner x \urcorner, \mathbf{n}) = \ulcorner A(\mathbf{n}) \urcorner$ ([2]) we define for formulas $A(x)$ with at most x free:

$$\Box[\ulcorner A(x) \urcorner] =_{\mathrm{def}} \Box(\mathrm{sub}(\ulcorner A(x) \urcorner, \ulcorner x \urcorner, x)).$$

If A has more than one free variable then we can iterate the use of sub and obtain similar definitions (all denoted by $\Box[.]$).

In what follows we write $\Box(A)$ and $\Box[A]$ for $\Box(\ulcorner A \urcorner)$ resp. $\Box[\ulcorner A \urcorner]$. Since for sentences A $\mathrm{PA} \vdash \Box[A] \leftrightarrow \Box(A)$ we will drop the notations $\Box(.)$ and $\Box[.]$ altogether and just use $\Box$ (probably using brackets for grouping) to mean $\Box[.]$[11].

Besides (26) we assume the Löb derivability conditions ([2]): for all formulas A, B

1. $\mathrm{PA} \vdash \Box A \to \Box\Box A$,
2. $\mathrm{PA} \vdash \Box(A \to B) \to (\Box A \to \Box B)$,
3. $\mathrm{PA} \vdash \Box(\Box A \to A) \to \Box A$.

Using the provability predicate and a predicate $\Sigma_1!(x)$ for syntactic Σ_1-ness we define

$$\Sigma_1(x) =_{\mathrm{def}} \exists \sigma \Sigma_1!(\sigma) \wedge \Box(x \dot{\leftrightarrow} \sigma).$$

And we define a binary predicate $\rhd$, expressing Π_1-conservativity, as follows:

$$x \rhd y =_{\mathrm{def}} \forall \pi (\Pi_1(\pi) \to (\Box(y \dot{\to} \pi) \to \Box(x \dot{\to} \pi))).$$

$\Pi_1(x)$ is a shorthand for $\Sigma_1(\dot{\neg} x)$.

Now we can interpret a ΣILM-formula in PA.

Definition 4.1 (Arithmetical realization $*$). An *arithmetical realization* $*$ is a function 'ΣILM-formulas' $\longrightarrow$ 'sentences of PA' which satisfies:

1. p^* is a sentence if $p \in \mathrm{PROP}$,
2. Commutativity with boolean connectives (e.g. $(A \wedge B)^* = (A^*) \wedge (B^*)$),
3. $(A \rhd B)^* = \ulcorner A^* \urcorner \rhd \ulcorner B^* \urcorner$,
4. $(\Sigma_1 A)^* = \Sigma_1(\ulcorner A^* \urcorner)$.

We say that a ΣILM-formula A is arithmetically valid if PA proves A^* for each arithmetical realization $*$.

In the case of PA we can change the notion of an arithmetical realization by changing the meaning of $A \rhd B$ to: '$\mathrm{PA} + A$ relatively interprets $\mathrm{PA} + B$'. In fact this notion coincides with the notion of Π_1-Conservativity [10]. However when considering Π_1-Conservativity (as we do) the results below do extend to certain finitely axiomatized theories (like for example $\mathrm{I}\Sigma_n$ for $n \geq 1$) but this is not the case if we change to relative interpretability.

Theorem 4.2. *If* $\Sigma\mathrm{ILM} \vdash A$ *and* $*$ *is an arithmetical realization then* $\mathrm{PA} \vdash A^*$.

[11] There is danger of confusion here since if $A(x)$ has x free then $\Box(\ulcorner A(x) \urcorner)$ means $\mathrm{PA} \vdash \forall x A(x)$ and $\Box[A(x)]$ means: for all $\mathbf{n}$ $\mathrm{PA} \vdash A(\mathbf{n})$. We take the risk.

Proof. Clearly the class of arithmetically valid formulas is closed under modus ponus and, by (26) under necessitation so it is sufficient to prove that all axiom schemas of ΣILM are arithmetically valid.

The validity of the GL part in ΣILM holds by choice of the predicate $\Box$ (for an explicit definition of such a predicate see ([2])). The validity of ΣL is therefore trivial. The validity of M(Σ) is discussed in Section 1.2 so what is left is the validity of the schemas 3-7 of Definition 1.3.

In the following let $*$ be some realization, reason in PA and let π be some Π_1 sentence.

3. Suppose $\Box(A^* \to B^*)$. If $\Box(B^* \to \pi)$ then $\Box(A^* \to \pi)$.

4. Suppose $A^* \rhd B^*$ and $B^* \rhd C^*$. If $\Box(C^* \to \pi)$ then $\Box(B^* \to \pi)$ and thus $\Box(A^* \to \pi)$.

5. Suppose $A^* \rhd C^*$ and $B^* \rhd C^*$. If $\Box(C^* \to \pi)$ then $\Box(A^* \to \pi)$ and $\Box(B^* \to \pi)$ and thus $\Box((A \vee B)^* \to \pi)$.

6. Suppose $A^* \rhd B^*$. If $\neg\Diamond B^*$ then $B^* \rhd \bot$ (trivially the equivalence $\Box\neg p \leftrightarrow p \rhd \bot$ is arithmetically valid) and thus $A^* \rhd \bot$ by Item 4 of this theorem.

7. Suppose $\Box(A^* \to \pi)$. Then $\Box(\Box\neg\pi \to \Box\neg A^*)$ and thus

$$\Box(\Diamond A^* \to \Diamond\pi).$$

π is Π_1 so $\Box(\neg\pi \to \Box\neg\pi)$ and thus:

$$\Box(\Diamond\pi \to \pi).$$

Conclusion: $\Box(\Diamond A^* \to \pi)$. $\dashv$

4.1 Arithmetical completeness of ΣILM

Now we are going to prove that all arithmetically valid ΣILM-formulas are theorems of ΣILM:

Theorem 4.3. *If A is a ΣILM-formula and $\Sigma\mathrm{ILM} \nvdash A$ then there exists some arithmetical realization $*$ such that $\mathrm{PA} \nvdash A^*$.*

The first theorem of this kind is the arithmetical completeness theorem for GL [17][2]. A lot of variations on this theorem are formulated and proved. One of them is the arithmetical completeness theorem for ILM (under the interpretation of Π_1-Conservativity) [9][10]. In the rest of this section we extend the proof as given in [10] by adding the Σ_1-case to obtain a proof of Theorem 4.3.

Let A be some ΣILM-formula not provable in ΣILM. Then by the extended modal completeness Theorem 3.17 there exist a strong Σ_1 ΣILM-model $M' = \langle W', R', S', V'\rangle$ and a world $w \in M'$ such that $w \models \neg A$. Without loss of generality assume $W' = \{1, 2, \ldots, \mathbf{n}\}$, $w = 1$ and w is a root of M': $\forall w' \in M' : w \neq w' \Rightarrow wR'w'$. We define a new ΣILM-model $M = \langle W, R, S, V\rangle$ as follows:

1. $W = W' + \{0\}$,
2. $R = R' + \{(0, w) \mid w \in W'\}$,

3. $S = S' + \{(0,x,y) \mid (1,x,y) \in S'\} + \{(0,1,x) \mid x \in W'\}$,

4. $V(p) = V'(p)$ if $1 \notin V'(p)$ and $V(p) = V'(p) + \{0\}$ otherwise.

Evidently we have for each $\mathbf{n} \geq 1$: $M, \mathbf{n} \models B \Leftrightarrow M', \mathbf{n} \models B$ and thus in particular:

$$M, 1 \models \neg A.$$

We are going to embed M into PA: assign sentences in the language of PA to the worlds of this model. The sentence assigned to a world w will we denote by $\lim_{\mathbf{w}}$. We will show then that it suffices to take the arithmetical realization $*$ such that $p^* = \bigvee_{v \in V(p)} \lim_{\mathbf{v}}$.

For $w \in M$ the sentence $\lim_{\mathbf{w}}$ expresses that a function $h : \omega \longrightarrow W$ (to be defined below) has $\mathbf{w}$ as a limit.

In order for PA to prove certain things about these sentences PA should be able to talk about M. This is accomplished by identifying the worlds with numbers, as we did above. To talk about the properties of M we define the following predicates:

$$R(x,y) =_{\text{def}} \bigvee_{uRv} (x{=}\mathbf{u} \wedge y{=}\mathbf{v}),$$
$$S(z,x,y) =_{\text{def}} \bigvee_{uS_tv} (z{=}\mathbf{t} \wedge x{=}\mathbf{u} \wedge y{=}\mathbf{v}).$$

In what follows we will, as we did in the modal case, write xS_zy for $S(z,x,y)$ and xRy for $R(x,y)$. Notice that all the ΣILM-frame properties are Δ_0 expressible and therefore, by Σ_1-completeness PA proves them.

For each $m \in \omega$ let $F_{\mathbf{m}}$ denote the formula with Gödel number $\mathbf{m}$. Let $\lambda mu.\lim(m,u)$ be the primitive recursive function such that for each $\mathbf{m}$ and $\mathbf{u}$:

$$\lim(\mathbf{m},\mathbf{u}) = \ulcorner \exists y (y{=}\mathbf{u} \wedge \exists w \forall x{\geq}w F_{\mathbf{m}}) \urcorner.$$

So if $\mathbf{f}$ is the code of a formula $F(x,y)$ that defines a function f, then $\lim(\mathbf{f},\mathbf{u})$ is the code of the statement: '$\mathbf{u}$ is the limit of f'.

Let $\lambda hwun.\Delta_h(w,u,n)$ be the primitive recursive function such that for all $\mathbf{h}$, $\mathbf{w}$, $\mathbf{u}$ and $\mathbf{n}$:

$$\Delta_{\mathbf{h}}(\mathbf{w},\mathbf{u},\mathbf{n}) = \ulcorner \exists t{>}\mathbf{n} \{(\exists y \exists x [y{=}\mathbf{u} \wedge x{=}t \wedge F_{\mathbf{h}}]) \wedge \forall x (\mathbf{n}{\leq}x{<}t \rightarrow \exists y (y{=}\mathbf{w} \wedge F_{\mathbf{h}}))\} \urcorner.$$

If $\mathbf{f}$ is the code of a formula $F(x,y)$ that defines a function f then $\Delta_{\mathbf{f}}(\mathbf{w},\mathbf{u},\mathbf{n})$ is the code of the statement: 'For some $t > \mathbf{n}$ $f(t) = \mathbf{u}$ and for each x: $\mathbf{n} \leq x < t$ implies $f(x) = \mathbf{w}$.'

Before we define our function h we need two more preliminary definitions. First it is well known that truth for Σ_1-sentences is definable([16]). There exists a Σ_1 formula Σ_1-Tr(x) such that:

$$\text{PA} \vdash \Sigma_1!(\sigma) \rightarrow \Box(\sigma \leftrightarrow \Sigma_1\text{-Tr}(\sigma)).$$

Secondly let RegWit(w, x) be the primitive recursive predicate for which

$$\text{for all } x\text{: } \mathbf{N} \models \Sigma_1\text{-Tr}(x) \Leftrightarrow \mathbf{N} \models \exists w \text{RegWit}(w, x).$$

Now define the function $\lambda hx.H_h(x)$ as follows:

1. $H_h(0) = 0$,
2. if $H_h(x)$ is defined then:
   (a) If for some $\mathbf{u}$: $H_h(x)R\mathbf{u}$ and for some $m \leq x$:
      i. $m \leq y < x \rightarrow H_h(y) = H_h(x)$ and
      ii. $\text{Pf}(x, \lim(h, \mathbf{u}) \dot{\rightarrow} \dot{\neg} \Delta_h(H_h(x), \mathbf{u}, m))$

      then $H_h(x + 1) = \mathbf{u}$,
   (b) else if for some m there exist $\mathbf{u}$, σ, z such that
      i. $H_h(x)S_{H_h(m)}\mathbf{u}$,
      ii. $m < x$ and $\text{Pf}(m, \lim(h, \mathbf{u}) \dot{\rightarrow} \dot{\neg} \sigma)$,
      iii. $\Sigma_1!(\sigma)$,
      iv. $z < x$ and RegWit(z, σ).

      If moreover for each $m' < m$ there do not exist $\mathbf{u}, \sigma, z$ with the above four properties, then $H_h(x + 1) = \mathbf{u}$,
   (c) in all other cases: $H_h(x + 1) = H_h(x)$.

The function $\lambda hx.H_h(x)$ is built up from primitive recursive case distinction and primitive recursion and is therefore primitive recursive. Let $H(h, x, y)$ be a Σ_1-formula defining this primitive recursive function and for which PA $\vdash$ $\forall hx \exists y H(h, x, y)$. Apply the diagonal lemma to find a formula $h(x, y)$ such that

$$\text{PA} \vdash h(x, y) \leftrightarrow H(\ulcorner h(x, y) \urcorner, x, y).$$

$h(x, y)$ defines a function which, of course, is primitive recursive. In what follows we will write $h(x) = y$ for $h(x, y)$.

Next define, for $\mathbf{u} \in \mathbf{W}$ the sentences $\lim_{\mathbf{u}}$:

$$\lim_{\mathbf{u}} =_{\text{def}} \exists y (y{=}\mathbf{u} \wedge \exists w \forall x{\geq}w h(x){=}y).$$

And the formula $\Delta_{wu}(n)$:

$$\Delta_{wu}(n) =_{\text{def}} \exists t{>}n\{(\exists y \exists x\{y{=}u \wedge x{=}t \wedge h(x){=}y\}) \wedge \forall x (n{\leq}x{<}t \rightarrow \exists y (y{=}w \wedge h(x){=}y))\}.$$

Now let us verify that in the definition of h (that is: the formula $H(\ulcorner h(x, y) \urcorner, x, y)$) we can replace all occurrences of $\lim(\ulcorner h(x, y) \urcorner, \mathbf{u})$ with $\ulcorner \lim_{\mathbf{u}} \urcorner$ to obtain a formula $H'(x, y)$ such that (still) PA $\vdash h(x, y) \leftrightarrow H'(x, y)$. For this it is sufficient that:

$$\text{PA} \vdash \lim(\ulcorner h(x, y) \urcorner, \mathbf{u}){=}\ulcorner \lim_{\mathbf{u}} \urcorner.$$

But this is clear since $\lambda xy.\mathrm{lim}(x,y)$ is primitive recursive and $\mathrm{lim}(\ulcorner h(x,y)\urcorner, \mathbf{u}) = \ulcorner \mathrm{lim}_{\mathbf{u}} \urcorner$ is true. Similarly one verifies that we can replace all occurrences of $\Delta_{\ulcorner h(x,y)\urcorner}(\mathbf{w}, \mathbf{u}, n)$ with $\Delta_{\mathbf{wu}}(n)$.

Before we move on let us agree on some terminology. If $h(x+1)$ is determined by Case 2a above then we say that h makes an R move at x. If $h(x+1)$ is determined by Case 2b then we say that h makes a S move at x. If in the latter case in addition $h(x) \neq h(x+1)$ then we say that h makes a *real* S move at x.

If h makes an S move then for some m there exist $\mathbf{u}, \sigma, z$ such that conditions 2(b)ii-2(b)iv are satisfied and for no $m' < m$ there exist such $\mathbf{u}, \sigma, z$. This m will be called the rank of the S move. Moreover we will drop the dots in $\dot{\neg}$, $\dot{\to}$,.... From the context it will be clear what is meant.

Lemma 4.4.

1. $\mathrm{PA} \vdash \bigvee_{w \in W} \mathrm{lim}_{\mathbf{w}}$,
2. $\mathbf{w} \neq \mathbf{u} \Rightarrow \mathrm{PA} \vdash \neg(\mathrm{lim}_{\mathbf{w}} \wedge \mathrm{lim}_{\mathbf{u}})$,
3. $\mathbf{w}R\mathbf{u} \Rightarrow \mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \neg\Box\neg\mathrm{lim}_{\mathbf{u}}$,
4. $\mathbf{w} \neq 0$ *and* $\neg\mathbf{w}R\mathbf{u} \Rightarrow \mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \Box\neg\mathrm{lim}_{\mathbf{u}}$,
5. *If* $\mathbf{u}S_{\mathbf{w}}\mathbf{v}$ *then*
$$\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \mathrm{lim}_{\mathbf{u}} \rhd \mathrm{lim}_{\mathbf{v}},$$
6. *If* $\mathbf{w} \neq 0$ *and* $M \subseteq \{w' \mid \mathbf{w}Rw'\}$ *such that for all* x, y, w'*:* $x \in M$, $\mathbf{w}Z \cdots Zw'$ *where* $Z = R \cup \bigcup_{\mathbf{t} \in W} S_{\mathbf{t}}$ *and* $xS_{w'}y$*, imply* $y \in M$ *then:*
$$\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \Sigma_1 \bigvee_{v \in M} \mathrm{lim}_{\mathbf{v}},$$
7. *If* $\mathbf{w}R\mathbf{u}$ *and* $V \subseteq W$ *such that for all* $\mathbf{v} \in V$ *we have not* $\mathbf{u}S_{\mathbf{w}}\mathbf{v}$*. Then*
$$\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \neg(\mathrm{lim}_{\mathbf{u}} \rhd \bigvee_{v \in V} \mathrm{lim}_{\mathbf{v}}).$$
8. *If* $\mathbf{u}S_{\mathbf{w}}\mathbf{v}$, $\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{u}} \to A$ *and* $\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{v}} \to \neg A$ *then*
$$\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \neg\Sigma_1 A,$$

Before we show Lemma 4.4 let us show some preliminary lemma's.

Lemma 4.5. *For any formula* $A = A(x,y)$*:*
$$\mathrm{PA} \vdash \forall x \Box(A(x,y)) \to \Box(\forall x{<}z A(x,y))$$

Proof. First note that $\mathrm{PA} \vdash \forall x{<}z A(x,y) \wedge A(z,y) \to \forall x{<}z{+}1 A(x,y)$. And therefore
$$\mathrm{PA} \vdash \Box(\forall x{<}z A(x,y) \wedge A(z,y) \to \forall x{<}z{+}1 A(x,y)). \tag{27}$$

Reason in PA. Induction on z. The case $z = 0$ is trivial. Assume $\forall x \Box(A(x,y)) \to \Box(\forall x{<}z A(x,y))$ and $\forall x \Box(A(x,y))$. Then both $\Box(\forall x{<}z A(x,y))$ and $\Box(A(z,y))$. Applying (27) we get $\Box(\forall x{<}z{+}1 A(x,y))$. ⊣

Lemma 4.6. (PA $\vdash$) *For any z and $\mathbf{w} \in W$. PA $+ \lim_{\mathbf{w}}$ proves that no S move to $\mathbf{w}$ can have rank less than z.*

Proof. Let

$$A(z,\lambda) = \exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda) \wedge \Sigma_1!\lambda \to \neg\Sigma_1\text{-}\mathrm{Tr}(\lambda).$$

We will show:

$$\mathrm{PA} \vdash \Box(\lim_{\mathbf{w}} \to \forall\lambda{<}zA(z,\lambda)). \tag{28}$$

Suppose

$$\mathrm{PA} \vdash \Box(\lim_{\mathbf{w}} \to (\exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda) \to \neg\lambda)). \tag{29}$$

Then

$$\mathrm{PA} \vdash \Box(\lim_{\mathbf{w}} \to (\exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda) \wedge \Sigma_1!(\lambda) \to \neg\Sigma_1\text{-}\mathrm{Tr}(\lambda))).$$

Substitution gives

$$\mathrm{PA} \vdash \Box(\lim_{\mathbf{w}} \to A(z,\lambda))$$

and thus (28) follows by Lemma 4.5.

So it remains to show (29). We have:

$$\begin{aligned}\mathrm{PA} \vdash \exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda) &\to \Box(\lim_{\mathbf{w}} \to \neg\lambda)\\ &\to \Box(\lim_{\mathbf{w}} \to (\exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda) \to \neg\lambda))\end{aligned}$$

and

$$\begin{aligned}\mathrm{PA} \vdash \neg\exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda) &\to \Box(\neg\exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda))\\ &\to \Box(\exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda) \to \neg\lambda)\\ &\to \Box(\lim_{\mathbf{w}} \to (\exists m{<}z\mathrm{Pf}(m, \lim_{\mathbf{w}} \to \neg\lambda) \to \neg\lambda)).\end{aligned}$$

From which (29) follows at once. $\dashv$

Lemma 4.7. PA $\vdash \forall xw(h(x)Rw \to \forall y{\leq}xh(y)Rw)$.

Proof. Reason in PA and suppose $h(x)Rw$. We show by induction on $n = x - y$ that $y{\leq}x \to h(y)Rw$. The case $n = 0$ is trivial. Suppose $h(y{+}1)Rw$. There are three cases to consider (1) $h(y){=}h(y{+}1)$, (2) $h(y)Rh(y{+}1)$ and (3) for some v: $h(y)S_vh(y{+}1)$. In case (1) there is nothing to prove. In case (2) $h(y)Rw$ by transitivity of R and in case (3) $h(y)Rw$ by the M property. $\dashv$

Lemma 4.8. (PA $\vdash$) *Consecutive real S moves have decreasing rank.*

Proof. Reason in PA. Assume: $\forall x(i{+}1{\leq}x{\leq}j \to h(x){=}h(j))$ and h makes a real S move at i and at j. Then there exist m, σ and z such that conditions 2(b)ii-2(b)iv of the definition of h are satisfied with $u = h(i+1)$ and $x = i$ and m the rank of this S move at i. But then, since $j{>}i$ and $h(j){=}h(i+1)$, m, σ and z satisfy these conditions also for $u = h(j)$ and $x = j$ and thus since $h(j+1){\neq}h(j)$ the rank of the S move at j must be $< m$. $\dashv$

Proof of Lemma 4.4. 1. First we show by induction on the converse of R that for each $\mathbf{w} \in W$:

$$\mathrm{PA} \vdash h(x){=}\mathbf{w} \rightarrow \exists z{\geq}x \forall y{\geq}z \neg h(y) R h(y{+}1). \tag{30}$$

The (IH) yields:

$$\mathrm{PA} \vdash \bigvee_{wRu} h(x){=}\mathbf{u} \rightarrow \exists z{\geq}x \forall y{\geq}z \neg h(y) R h(y{+}1). \tag{31}$$

Now reason in PA and assume $h(x){=}\mathbf{w}$. If there exists some z such that $x{\leq}z$ and $h(z)Rh(z{+}1)$ then for this z we have, by Lemma 4.7, $h(x){=}\mathbf{w}Rh(z{+}1)$, that is $\bigvee_{wRu} h(z{+}1){=}\mathbf{u}$, and thus (30) follows from (31). In case such a z does not exist (30) follows at once.

Taking $\mathbf{w} = 0$ and $x = 0$ in (30) we obtain:

$$\mathrm{PA} \vdash \exists z \forall y{\geq}z \neg h(y) R h(y{+}1). \tag{32}$$

In other words: PA proves that after some point h does not make any R moves.

Now reason in PA. Pick z such that after z there are no more R moves. If there are no more real S moves either then $h(z)$ is the limit of h. So suppose h does make some real S move after z. Let h make a real S move at $i \geq z$ with minimal rank. In other words pick a minimal m such that for some $\mathbf{u}$ there exists $i \geq z$:

$$\exists \sigma{<}m \exists w{<}i (h(i) S_{h(m)} \mathbf{u} \wedge \mathrm{Pf}(m, \lim_{\mathbf{u}} \rightarrow \neg\sigma) \wedge \Sigma_1!(\sigma) \wedge \mathrm{RegWit}(w, \sigma)). \tag{33}$$

Fix such i and $\mathbf{u}$. By Lemma 4.8 any real S move after i must have rank lower than m which by the minimality of m is impossible, conclusion: $h(i+1)$ (that is $\mathbf{u}$) is the limit of h.

2. h is primitive recursive so PA proves $\exists! y h(x){=}y$. So it follows at once that h cannot have two different limits.

3. We have:

$$\mathrm{PA} \vdash \Box\neg\lim_{\mathbf{u}} \rightarrow \Box(\lim_{\mathbf{u}} \rightarrow \neg\Delta_{\mathbf{wu}}(z)),$$

so

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \wedge \Box\neg\lim_{\mathbf{u}} \rightarrow \exists z (\forall z'{\geq}z h(z'){=}\mathbf{w} \wedge \Box(\lim_{\mathbf{u}} \rightarrow \neg\Delta_{\mathbf{wu}}(z))).$$

Since PA proves that any provable sentence has arbitrary long proofs:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \wedge \Box\neg\lim_{\mathbf{u}} \rightarrow \exists z \exists m{\geq}z ((\forall z'{\geq}z h(z'){=}\mathbf{w}) \wedge \mathrm{Pf}(m, \lim_{\mathbf{u}} \rightarrow \neg\Delta_{\mathbf{wu}}(z))).$$

But

$$\mathrm{PA} \vdash (\exists z \{m{\geq}z \wedge \forall z'{\geq}z h(z'){=}\mathbf{w} \wedge \mathrm{Pf}(m, \lim_{\mathbf{u}} \rightarrow \neg\Delta_{\mathbf{wu}}(z))\}) \rightarrow h(m{+}1){=}\mathbf{u} \wedge h(m{+}1){=}\mathbf{w},$$

By assumption: $\mathbf{w}R\mathbf{u}$ so $\mathbf{w} \neq \mathbf{u}$ and $\mathrm{PA} \vdash \mathbf{u}{\neq}\mathbf{w}$ and thus:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \wedge \Box\neg\lim_{\mathbf{u}} \rightarrow \bot.$$

4. First assume $\mathbf{w} = \mathbf{u} = \mathbf{n} \neq 0$. In this case the statement of the lemma boils down to:

$$\text{PA} \vdash \lim_{\mathbf{n}} \to \Box\neg\lim_{\mathbf{n}}.$$

Since $\mathbf{n} \neq 0$ and $\text{PA} \vdash h(0){=}0$: $\text{PA} \vdash \exists x h(x){=}\mathbf{n} \to \exists x(h(x){\neq}\mathbf{n} \wedge h(x{+}1){=}\mathbf{n})$. So since $\text{PA} \vdash \lim_{\mathbf{n}} \to \exists x h(x, \mathbf{n})$ we have:

$$\text{PA} \vdash \lim_{\mathbf{n}} \to \exists x(h(x){\neq}\mathbf{n} \wedge h(x{+}1){=}\mathbf{n}).$$

By definition of h:

$$\begin{aligned}\text{PA} \vdash (h(x){\neq}\mathbf{n} \wedge h(x{+}1){=}\mathbf{n}) \to &\exists y\{\forall z[y{\leq}z{\leq}x \to h(z){=}h(y)] \wedge \Box(\lim_{\mathbf{n}} \to \neg\Delta_{h(y)\mathbf{n}}(y))\} \\ &\vee \exists\sigma(\Sigma_1!(\sigma) \wedge \Sigma_1\text{-Tr}(\sigma) \wedge \Box(\lim_{\mathbf{n}} \to \neg\sigma)).\end{aligned}$$

By definition of Δ and by Σ_1 completeness:

$$\begin{aligned}\text{PA} \vdash (\forall z\{y{\leq}z{\leq}x \to h(z){=}h(y)\} \wedge h(x{+}1){=}\mathbf{n}) \to \Delta_{h(y)\mathbf{n}}(y)& \\ \to \Box(\Delta_{h(y)\mathbf{n}}(y)).&\end{aligned}$$

Moreover $\text{PA} \vdash \Sigma_1!(\sigma) \wedge \Sigma_1\text{-Tr}(\sigma) \to \Box(\sigma)$, so

$$\text{PA} \vdash \exists x(h(x){\neq}\mathbf{n} \wedge h(x{+}1){=}\mathbf{n}) \to \Box(\neg\lim_{\mathbf{n}}).$$

Now assume $\mathbf{w} \neq \mathbf{u}$. It is sufficient to show

$$\text{PA} \vdash h(x){=}\mathbf{w} \to \Box(\neg\lim_{\mathbf{u}}) \tag{34}$$

and since $\text{PA} \vdash h(x){=}\mathbf{w} \to \Box(h(x){=}\mathbf{w})$ in order to show (34) it suffices to show that

$$\text{PA} \vdash \Box(h(x){=}\mathbf{w} \to \neg\lim_{\mathbf{u}}). \tag{35}$$

We show that $\Box(h(x){=}\mathbf{w} \to \neg\lim_{\mathbf{u}})$ is true. Our argument can be formalized in PA to show (35).

So let $\mathbf{x} \in \omega$ and reason in PA. Suppose $h(\mathbf{x}){=}\mathbf{w}$ and assume for a contradiction that $\mathbf{u}$ is the limit of h. Since $\mathbf{u}{\neq}\mathbf{w}$ there must exists some $z{\geq}\mathbf{x}$ such that $h(z){\neq}\mathbf{u}$ and $h(z{+}1){=}\mathbf{u}$. Since not $h(\mathbf{x}){=}\mathbf{w}R\mathbf{u}{=}h(z{+}1)$ by Lemma 4.7:

$$\mathbf{x} \leq y \Rightarrow \text{ not } h(y)Rh(z{+}1). \tag{36}$$

So in particular $\neg h(z)Rh(z{+}1)$ and thus there exist some m and σ:

$$h(z)S_{h(m)}h(z{+}1) \wedge \text{Pf}(m, \lim_{\mathbf{u}} \to \neg\sigma) \wedge \Sigma_1!(\sigma) \wedge \Sigma_1\text{-Tr}(\sigma).$$

By Lemma 4.6 this m must be $\geq\mathbf{x}$. But since $h(z)S_{h(m)}h(z{+}1)$ we must have $h(m)Rh(z{+}1)$ and thus by (36) $m{<}\mathbf{x}$. A contradiction.

5. If $\mathbf{u} = \mathbf{v}$ the statement is trivial so assume $\mathbf{u} \neq \mathbf{v}$. Suppose

$$\text{PA} \vdash \Box\{h(z){=}\mathbf{w} \wedge \text{Pf}(z, \lim_{\mathbf{v}} \to \pi) \wedge \Pi_1!(\pi)\} \to \Box(\lim_{\mathbf{u}} \to \pi). \tag{37}$$

Reason in PA and assume $\lim_{\mathbf{w}}$. Choose π such that $\Pi_1(\pi)$ and $\Box(\lim_{\mathbf{v}} \to \pi)$. $\Pi_1(\pi)$ means that for some π': $\Pi_1!(\pi')$ and $\Box(\pi \leftrightarrow \pi')$ and thus w.l.o.g. we can assume $\Pi_1!(\pi)$. Choose z such that $\forall z'{\geq}z h(z'){=}\mathbf{w}$ and $\text{Pf}(z, \lim_{\mathbf{v}} \to \pi)$. Then:

a. $\Box(h(z){=}\mathbf{w})$ and

b. $\Box(\mathrm{Pf}(z, \mathrm{lim}_{\mathbf{v}} \to \pi))$.

Moreover since $\Pi_1!(\pi)$ is Δ_0:

c. $\Box(\Pi_1!(\pi))$.

Applying this to (37) gives: $\Box(\mathrm{lim}_{\mathbf{u}} \to \pi)$ so it is sufficient to show (37).

We will show that $\Box\{h(z){=}\mathbf{w} \wedge \mathrm{Pf}(z, \mathrm{lim}_{\mathbf{v}} \to \pi) \wedge \Pi_1(\pi)\} \to \Box(\mathrm{lim}_{\mathbf{u}} \to \pi)$ is true. Our argument can be formalized to show it provable.

So let $\mathbf{z}$, $\ulcorner\pi\urcorner \in \omega$ and assume

$$\Box(h(\mathbf{z}){=}\mathbf{w} \wedge \mathrm{Pf}(\mathbf{z}, \mathrm{lim}_{\mathbf{v}} \to \ulcorner\pi\urcorner) \wedge \Pi_1(\ulcorner\pi\urcorner)). \tag{38}$$

Reason in PA and assume $\mathrm{lim}_{\mathbf{u}}$. Assume for a contradiction $\neg\pi$. Since $\Sigma_1(\neg\ulcorner\pi\urcorner)$ there exists some w such that $\mathrm{RegWit}(w, \neg\ulcorner\pi\urcorner)$. Now choose $x \geq w, \mathbf{z}$ such that $h(x{+}1) = h(x) = \mathbf{u}$. This is possible since $\mathbf{u}$ is the limit of h. So in particular h does not make an R move at x. However, by (38) h does make an S move at x and the rank of this move is $\leq\mathbf{z}$. Now if the rank is $<\mathbf{z}$ then by Lemma 4.6 and since we have assumed $\mathrm{lim}_{\mathbf{u}}$, $h(x+1) \neq \mathbf{u}$. A contradiction. So the rank must be equal to $\mathbf{z}$. But this implies $h(x+1) = \mathbf{v} \neq \mathbf{u}$. Again a contradiction.

6. Let $\mathbf{w} \in W$, $\mathbf{w} \neq 0$ and $M \subset W$ as stated in the hypothesis. It is sufficient to show:

$$\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \Box(\bigvee_{u \in M} \exists x h(x){=}\mathbf{u} \to \bigvee_{u \in M} \mathrm{lim}_{\mathbf{u}}). \tag{39}$$

Assume:

$$\mathrm{PA} \vdash \Box(h(n){=}\mathbf{w} \wedge \bigvee_{m_0 \in M} \exists x h(x){=}\mathbf{m}_0 \to \bigwedge_{n_0 \notin M, wRn_0} \neg\mathrm{lim}_{\mathbf{n}_0}). \tag{40}$$

Combining Item 1 and Item 4 of this lemma:

$$\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \Box(\bigvee_{wRv} \mathrm{lim}_{\mathbf{v}}).$$

So, again using Item 1, (40) yields:

$$\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \Box(h(n){=}\mathbf{w} \wedge \bigvee_{m_0 \in M} \exists x h(x){=}\mathbf{m}_0 \to \bigvee_{n_0 \in M} \mathrm{lim}_{\mathbf{n}_0}).$$

Now (39) follows since $\mathrm{PA} \vdash \mathrm{lim}_{\mathbf{w}} \to \exists n h(n){=}\mathbf{w}$ and $\mathrm{PA} \vdash h(n){=}\mathbf{w} \to \Box(h(n){=}\mathbf{w})$.

So we have to show (40). Again, we will show that

$$\Box(h(n){=}\mathbf{w} \wedge \bigvee_{m_0 \in M} \exists x h(x){=}\mathbf{m}_0 \to \bigwedge_{n_0 \notin M, wRn_0} \neg\mathrm{lim}_{\mathbf{n}_0})$$

is true. Formalizing our argument will show it provable.

So let $\mathbf{n}$ be some number and reason in PA. Suppose $h(\mathbf{n})=\mathbf{w}$. Let $\mathbf{m}_0 \in M$ and assume $\exists x h(x)=\mathbf{m}_0$. Fix some x_0 such that $h(x_0) = \mathbf{m}_0$ and assume for a contradiction that $\lim_{\mathbf{n}_0}$ for some $\mathbf{n}_0 \notin M$, $\mathbf{w}R\mathbf{n}_0$. Pick $x_1 > x_0$ such that $h(x_1) = \mathbf{n}_0$. Let $u_0, u_1, \ldots, u_{k-1}, u_k$ be the sequence of all the values h assumes between x_0 and x_1. Let $t_0, t_1, \ldots, t_{k-1}$ be the sequence of values such that h makes a move from u_i towards u_{i+1} at t_i. Let $j > 0$ be minimal with property

$$\text{for all } i \text{ with } j \leq i \leq k \text{ not } u_0 R u_j. \tag{41}$$

Such j exists since $j = k$ satisfies this property.

By Lemma 4.7 we see that for these i not $u_{i-1}Ru_i$ and thus the moves at t_{i-1} for $j \leq i \leq k$ must all be S moves. Using 4.6 the rank of the move to $u_k = \mathbf{n}_0$ must be $\geq n$ and since consecutive S moves have decreasing rank (Lemma 4.8) this holds for all the moves between u_{j-1} and u_k and thus for some $n_j, n_{j+1}, \ldots, n_k \geq \mathbf{n}$:

$$j \leq i \leq k \Rightarrow \quad u_{i-1} S_{h(n_i)} u_i. \tag{42}$$

Obviously for each n_i there exist $c_0, c_1, \ldots, c_r$ s.t. $h(\mathbf{n}) = c_0 Z c_1 Z \cdots Z c_r = h(n_i)$ where $Z = R \cup \bigcup_{t \in W} S_t$. Moreover $u_0 = u_{j-1}$ or, by (41), $u_0 R u_{j-1}$ and in either case since $h(\mathbf{n}) R u_0$:

$$u_0 S_{h(\mathbf{n})} u_{j-1}. \tag{43}$$

Combining (42),(43) we can prove, with induction on i and using that $h(\mathbf{n}) = \mathbf{w}$, that each $u_i \in M$. But $u_k = \mathbf{n}_0 \notin M$. A contradiction.

7. Here we pay the price for our 'ugly' modal ΣILM semantics. If we (could) assume that M is a ΣL-model then we could derive this item from item 6. Now we have to copy large parts of the proof given there.

Assume

$$\text{for each } \mathbf{v} \in V\text{: PA} \vdash \Box(\lim_{\mathbf{v}} \to \neg\Delta_{\mathbf{wu}}(n)). \tag{44}$$

Notice: PA $\vdash \Pi_1(\neg\Delta_{\mathbf{wu}}(n))$. Now reason in PA and assume $\lim_{\mathbf{w}}$. Pick n such that for all $n' > n$, $h(n') = \mathbf{w}$. Then

$$\neg\Box(\lim_{\mathbf{v}} \to \neg\Delta_{\mathbf{wu}}(n)). \tag{45}$$

Since otherwise for some $n' > n$, $\mathrm{Pf}(n', \lim_{\mathbf{u}} \to \neg\Delta_{\mathbf{wu}}(n))$ and thus by definition of h: $h(n'{+}1) = \mathbf{u} \neq \mathbf{w}$, a contradiction. (44) implies

$$\Box(\bigvee_{v \in V} \lim_{\mathbf{v}} \to \neg\Delta_{\mathbf{wu}}(n)). \tag{46}$$

Combining (45) and (46) we conclude $\neg(\lim_{\mathbf{u}} \rhd \bigvee_{\mathbf{v} \in V} \lim_{\mathbf{v}})$. So we are left to show (44).

For each $\mathbf{v} \in V$ we show that $\Box(\lim_{\mathbf{v}} \to \neg\Delta_{\mathbf{wu}}(n))$ is true. Our argument can be formalized to show (44).

Pick $\mathbf{n} \in \omega$ and reason in PA. Assume $\lim_{\mathbf{v}}$ and assume for a contradiction that $\Delta_{\mathbf{wu}}(\mathbf{n})$. In other words:

$$\exists s{>}\mathbf{n}(h(s){=}\mathbf{u} \wedge \forall x(\mathbf{n} \leq x{<}s \to h(x){=}\mathbf{w})). \tag{47}$$

Fix such an s and choose $t > s$ such that for all $t' \geq t$: $h(t') = \mathbf{v}$. Let $u_0, u_1, \ldots, u_{k-1}, u_k$ be all the values h assumes between s and t. Let $t_0, t_1, \ldots, t_{k-1}$ be the sequence of values for which h make a move at t_i from u_i towards u_{i+1} and let $t_k = t$. Choose $j > 0$ minimal with the property

$$\text{for all } i \text{ with } j \leq i \leq k \text{ not } u_0 R u_j.$$

Then by Lemma 4.7 for each $j \leq i \leq k$: not $u_{i-1} R u_i$ and thus the moves at these t_{i-1} are S moves. If m_{i-1} is the rank of the S move at t_{i-1} then by definition of h: $m_{i-1} < t_{i-1}$ and since $h(m_{i-1}) R h(t_i) = u_i$ consequently: $m_{i-1} < s$. Moreover just like in Item 6 of this lemma these ranks are $\geq \mathbf{n}$ and thus by (47):

$$\text{for each } i\text{: } j \leq i \leq k \Rightarrow u_{i-1} S_{\mathbf{w}} u_i.$$

By choice of j: $u_0 R u_{j-1}$ or $u_0 = u_{j-1}$. In either case since $\mathbf{w} R u_0 (= \mathbf{u})$:

$$u_0 S_{\mathbf{w}} u_{j-1}.$$

By transitivity of $S_{\mathbf{w}}$ we conclude: $\mathbf{u} = u_0 S_{\mathbf{w}} u_k = \mathbf{v}$, a contradiction.

8. Let $\mathbf{u} S_{\mathbf{w}} \mathbf{v}$. By Item 5 of this lemma:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to (\lim_{\mathbf{u}} \rhd \lim_{\mathbf{v}}).$$

In other words:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to (\forall x\{\Box(\lim_{\mathbf{v}} \to x) \wedge \Box(\lim_{\mathbf{u}} \to \neg x) \to \neg \Pi_1(x)\}).$$

Now if $\mathrm{PA} \vdash \lim_{\mathbf{u}} \to A$ and $\mathrm{PA} \vdash \lim_{\mathbf{v}} \to \neg A$ then $\mathrm{PA} \vdash \Box(\lim_{\mathbf{u}} \to A)$ and $\mathrm{PA} \vdash \Box(\lim_{\mathbf{v}} \to \neg A)$ and thus $\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \neg \Pi_1(\ulcorner \neg A \urcorner)$. In other words:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \neg \Sigma_1(\ulcorner A \urcorner).$$

$\dashv$

In what follows we let $*$ be the realization such that $p^* = \bigvee_{v \in V(p)} \lim_{\mathbf{v}}$.

Lemma 4.9. *For each* ΣILM*-formula* A *and* $\mathbf{w} \in W$, $\mathbf{w} \neq 0$*:*

1. $\mathbf{w} \models A \Rightarrow \mathrm{PA} \vdash \lim_{\mathbf{w}} \to A^*$,
2. $\mathbf{w} \not\models A \Rightarrow \mathrm{PA} \vdash \lim_{\mathbf{w}} \to \neg A^*$.

Proof. Induction on the complexity of A.

If A is atomic then 1 is clear. 2 holds since PA proves that h cannot have two different limits.

Truth functional cases are trivial.

Case $A = \Sigma_1 A'$. Put $M = \{v \in W \mid \mathbf{w}Rv, v \models A'\}$. By the (IH): $\mathrm{PA} \vdash \bigvee_{v \in M} \lim_{\mathbf{v}} \to A'^*$ so

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box(\bigvee_{v \in M} \lim_{\mathbf{v}} \to A'^*). \tag{48}$$

By the (IH) $\mathrm{PA} \vdash \bigvee_{wRv, v \not\models A'} \lim_v \to \neg A'^*$ so

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box(\bigvee_{wRv, v \not\models A'} \lim_v \to \neg A'^*)$$

and thus since $\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box \bigvee_{wRv} \lim_{\mathbf{v}}$:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box(A'^* \to \bigvee_{v \in M} \lim_{\mathbf{v}}). \tag{49}$$

Combining (48) and (49) we get

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to (\Sigma_1 A'^* \leftrightarrow \Sigma_1 \bigvee_{v \in M} \lim_{\mathbf{v}}). \tag{50}$$

Suppose $\mathbf{w} \models \Sigma_1 A'$. M is closed under R and S_t steps, for t such that $\mathbf{w}(R \cup \bigcup_{v \in W} S_{\mathbf{v}})^* t$, so Lemma 4.4-6 gives $\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Sigma_1 \bigvee_{\mathbf{v} \in M} \lim_{\mathbf{v}}$ and thus (50) yields 1. Now suppose $\mathbf{w} \not\models \Sigma_1 A'$. Pick $\mathbf{u}$ and $\mathbf{v}$ such that $\mathbf{w}R\mathbf{u}S_{\mathbf{w}}\mathbf{v}$, $\mathbf{u} \models A'$ and $\mathbf{v} \not\models A'$ (these exist since M is strong Σ_1). Then by the (IH) $\mathrm{PA} \vdash \lim_{\mathbf{u}} \to A'^*$ and $\mathrm{PA} \vdash \lim_{\mathbf{v}} \to \neg A'^*$. Applying Lemma 4.4-8 we get 2.

Case $A = B \rhd C$. Suppose $\mathbf{w} \models A$. Put

$$\begin{aligned} V'_B =& \{v \mid \mathbf{w}Rv \text{ and } v \not\models B\}, \\ V_B =& \{v \mid \mathbf{w}Rv \text{ and } v \models B\} \\ V_C =& \{v \mid \mathbf{w}Rv \text{ and } v \models C\}. \end{aligned}$$

Then by (IH) $\mathrm{PA} \vdash B \to \neg \bigvee_{\mathbf{v} \in V'_B} \lim_{\mathbf{v}}$ and thus since $\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box \bigvee_{\mathbf{w}R\mathbf{v}} \lim_{\mathbf{v}}$:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box(B \to \bigvee_{v \in V_B} \lim_{\mathbf{v}}). \tag{51}$$

For each $\mathbf{v} \in V_B$ there exists $\mathbf{u}$ in V_C such that $\mathbf{v}S_{\mathbf{w}}\mathbf{u}$. By Lemma 5 for such $\mathbf{u}$: $\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \lim_{\mathbf{v}} \rhd \lim_{\mathbf{u}}$ so

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \bigvee_{v \in V_B} \lim_{\mathbf{v}} \rhd \bigvee_{u \in V_C} \lim_{\mathbf{u}}. \tag{52}$$

By (IH) $\mathrm{PA} \vdash \bigvee_{\mathbf{u} \in V_C} \lim_{\mathbf{u}} \to C^*$ and thus:

$$\mathrm{PA} \vdash \lim_w \to \Box(\bigvee_{u \in V_C} \lim_{\mathbf{u}} \to C^*). \tag{53}$$

Combining (51) (52) and (53): $\mathrm{PA} \vdash \lim_{\mathbf{w}} \to B^* \rhd C^*$.

Suppose $\mathbf{w} \not\models A$. Pick $\mathbf{u}$ s.t. $\mathbf{w}R\mathbf{u}$, $\mathbf{u} \models B$ and for all $\mathbf{u}S_{\mathbf{w}}\mathbf{v}$: $\mathbf{v} \not\models C$. Put

$$\begin{aligned} V &= \{v \mid \mathbf{w}Rv \text{ and } v \models C\}, \\ V' &= \{v \mid \mathbf{w}Rv \text{ and } v \not\models C\}. \end{aligned}$$

Then $v \in V$ implies not $\mathbf{u}S_{\mathbf{w}}v$ and thus by Lemma 4.4-7:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \neg(\lim_{\mathbf{u}} \rhd \bigvee_{v \in V} \lim_{\mathbf{v}}). \tag{54}$$

By (IH) $\mathrm{PA} \vdash \lim_{\mathbf{u}} \to B^*$ and thus

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box(\lim_{\mathbf{u}} \to B^*). \tag{55}$$

Since $\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box(\bigvee_{v \in V'} \lim_{\mathbf{v}} \vee \bigvee_{v \in V} \lim_{\mathbf{v}})$ and by (IH) $\mathrm{PA} \vdash C^* \to \neg \bigvee_{v \in V'} \lim_{\mathbf{v}}$ we have:

$$\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \Box(C^* \to \bigvee_{v \in V} \lim_{\mathbf{v}}). \tag{56}$$

Combining (54) (55) and (56) we conclude: $\mathrm{PA} \vdash \lim_{\mathbf{w}} \to \neg(B^* \rhd C^*)$. ⊣

Lemma 4.10. $\lim_0$ *is true.*

Proof. By Lemma 4.4-1 some $\lim_{\mathbf{v}}$ is true. If $\mathbf{v} \neq 0$ and $\lim_{\mathbf{v}}$ is true then by Lemma 4.4-4: $\Box\neg\lim_{\mathbf{v}}$, a contradiction. Conclusion: $\lim_0$ is true. ⊣

Now we're in position to finish the proof of Theorem 4.3.

Proof of Theorem 4.3. Since $1 \not\models A$: $\mathrm{PA} \vdash \lim_1 \to \neg A^*$ and thus

$$\mathrm{PA} \vdash \neg\Box\neg\lim_1 \to \neg\Box A^*.$$

By Item 3. of Lemma 4.4: $\mathrm{PA} \vdash \lim_0 \to \neg\Box\neg\lim_1$ so as $\lim_0$ is true (and PA is sound) $\neg\Box A^*$ is true as well. ⊣

4.2 ΣL

One can look at ΣILM as if it is an extension of ΣL and if one does so it makes sense to ask whether a ΣILM-formula A not containing $\rhd$ is provable in ΣL, is valid on a ΣL-frame, and so on.

The next theorems use this to derive the arithmetical completeness theorem for ΣL from the arithmetical completeness theorem for ΣILM.

Theorem 4.11 (Conservativity). *If A is a ΣL-formula (in other words: a ΣILM-formula that does not contain $\rhd$) and $\Sigma\mathrm{ILM} \vdash A$ then $\Sigma\mathrm{L} \vdash A$.*

Proof. Let A be as in the hypothesis of the theorem. Then by the modal soundness Theorem 3.12 for ΣILM: A valid on all ΣILM-frames (using the ΣILM forcing relation). But each ΣL-frame is equivalent (in the sense that the same ΣL-formulas are valid on them, see Theorem 2.25) to a ΣILM-frame with the additional properties: for all x, y, a, b, t

1. xRy and aS_yb implies aS_xb,
2. xS_ty and aS_yb implies aS_xb.

One easily proves, with induction on the complexity of A, that A is valid on all ΣILM-frames using the forcing relation for ΣILM-formulas iff A is valid on all ΣILM-frames satisfying 1 and 2 using the ΣL forcing relation. Applying the modal completeness Theorem 2.8 for ΣL gives: $\Sigma\mathrm{L} \vdash A$. $\dashv$

Theorem 4.12 (Arithmetical completeness). *Let A be a ΣL-formula. If $\Sigma\mathrm{L} \nvdash A$ then there exists an arithmetical interpretation $*$ such that $\mathrm{PA} \nvdash A^*$.*

Proof. If A is a ΣL-formula and $\Sigma\mathrm{L} \nvdash A$ then by Theorem 4.11 $\Sigma\mathrm{ILM} \nvdash A$ and thus by the arithmetical completeness theorem for ΣILM there exists some arithmetical interpretation $*$ such that $\mathrm{PA} \nvdash A^*$. $\dashv$

5 Fragments and variations

Above we have studied formulas which are always provable (e.a. formulas A for which for any realization $*$: $\mathrm{PA} \vdash A^*$). A variation could be to study the formulas A that are always true: for any realization $*$: $\mathbf{N} \models A^*$. As is the case with ILM the addition of reflxection (the principle $\Box A \to A$) will probably do.

Another subject of study could be the determination of certain fragments of the above logics. For instance the formulas that do only contain unnested occurrences of the operator Σ_1.

What follows is a copy of a note from J. Joosten [11] which can be categorized under these subject matters. The always true formulas of the form $\bigwedge_{0\leq i\leq n} \Sigma_1 B_i \to \Sigma_1 A$, where A and all the B_i's are propositional, are determined.

5.1 The propositional Σ_1-logic of PA

In this section three consequence relations will be introduced. These consequence relations will concern the language of modal logic. First we will restrict ourselves to the propositional case only. The main result will be to proof that the three relations are the same.

5.1.1 An arithmetical consequence relation $\models_{\Sigma_1}$

Definition 5.1. An *arithmetical translation* is a function $*$ assigning arithmetical sentences to propositional formulas (in our case) in such a way that

$(\bot)^* = \bot$, and $*$ "distributing over" the connectives, in the sense that, for example, $(p \wedge q)^* = p^* \wedge q^*$.

In the following definition the $*$ is suppose to range over all possible arithmetical translations.

Definition 5.2. $\Delta \models_{\Sigma_1} \phi \Leftrightarrow \forall * [\forall \delta \in \Delta \;\; \delta^* \in \Sigma_1(\mathrm{PA}) \to \phi^* \in \Sigma_1(\mathrm{PA})]$

5.1.2 The syntactical consequence relation $\vdash_{\Sigma_1}$

Definition 5.3. $\vdash_{\Sigma_1}$is the binary relation between sets of propositional formulas and propositional formulas such that $\Gamma \vdash_{\Sigma_1} \phi$ iff some conjunction of disjunctions of formulas in Γ is equivalent to ϕ.

Instead of $\Gamma \cup \{\psi\} \vdash_{\Sigma_1} \phi$ we will often write $\Gamma, \psi \vdash_{\Sigma_1} \phi$.

One can give a more deduction-like definition of $\vdash_{\Sigma_1}$ (with for example a Cut rule, a Weakening rule etc.). In such a formulation it is then clear that $\vdash_{\Sigma_1}$ is a consequence relation. The above definition is in the present context more useful and informative.

5.1.3 The semantical consequence relation $\Vdash_{\Sigma_1}$

Where does the intuition come from and how to depict it.

Definition 5.4. An *S-model* is a pair $< \{l, r\}, \Vdash >$ where $\Vdash$ is a forcing relation telling which propositional variables are forced on l and r i.e. $\Vdash : \{l, r\} \to Prop$ and $l \Vdash \top$ and $r \Vdash \top$.

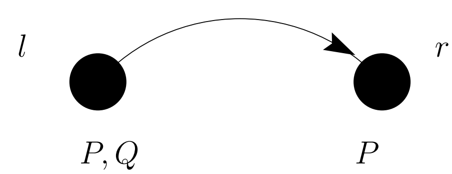
The forcing relation extends on the natural way to sentences by stating that it "distributes over the logical connectives". So for example $l \Vdash p \wedge q \Leftrightarrow l \Vdash p \;\&\; l \Vdash q$. The letters l and r stands for 'left' and 'right' respectively as we depict them correspondingly.

Definition 5.5. Let M denote some S-model. The relation $\models\!\!\!>$is defined by $M \models\!\!\!> \phi \Leftrightarrow (l \Vdash \phi \to r \Vdash \phi)$ and $M \models\!\!\!> \Gamma \Leftrightarrow \forall \gamma \in \Gamma \;\; M \models\!\!\!> \gamma$.

In this definition of $\models\!\!\!>$we use persistence of forcing from left to right. To indicate this direction in our drawings we will connect l to r with an arrow.

Definition 5.6. $\Gamma \Vdash_{\Sigma_1} \phi \Leftrightarrow$ for all S-models M $[M \models\!\!\!> \Gamma \to M \models\!\!\!> \phi]$.

We write $\Gamma \not\models\!\!\!> \phi$ for $\neg(\Gamma \models\!\!\!> \phi)$. Likewise we write $\Gamma \nVdash_{\Sigma_1} \phi$ for $\neg(\Gamma \Vdash_{\Sigma_1} \phi)$. In the latter case we can find a model witnessing this fact. So, for example $P \nVdash_{\Sigma_1} P \wedge Q$ is demonstrated in figure 1; an S-model.



For indeed $M \models\!\!> P$ but $M \not\models\!\!> P \wedge Q$. In a picture we only indicate which propositional letters are forced. Those letters not mentioned are not forced.

Theorem 5.7. *The three above defined consequence relations, $\models_{\Sigma_1}$, $\vdash_{\Sigma_1}$, and $\Vdash_{\Sigma_1}$ coincide.*

Proof. The Proof will consist of three steps.

A.) $\vdash_{\Sigma_1} \subseteq \ \models_{\Sigma_1}$; This is trivial and does not need a proof.

B.) $\models_{\Sigma_1} \subseteq \ \Vdash_{\Sigma_1}$; We reason by contraposition. So, suppose that $\Gamma \not\Vdash_{\Sigma_1} \phi$. In this case we can find an S-model M with $M \models\!\!> \Gamma$ and $M \not\models\!\!> \phi$. For the propositional variables we can distinguish four different situations depending on the variables being forced on l and r or not. We now define a map $h : Prop \longrightarrow \{\bot, \top, \Box\bot, \Diamond\top\}$ as disposed in the table below.

	l	r	$h(p)$
p	$+$	$+$	$\top$
p	$+$	$-$	$\Diamond\top$
p	$-$	$+$	$\Box\bot$
p	$-$	$-$	$\bot$

So, for example, if $l \not\Vdash p$ and $r \Vdash p$, $h(p)$ will take the value of $\Box\bot$. We see that $h(p) = \Diamond\top \Leftrightarrow M \not\models\!\!> p$. Now $\{\bot, \top, \Box\bot, \Diamond\top\}$ can be made into a Boolean algebra in the obvious way by considering it as a subalgebra of the Magari algebra. Thus we can extend h to $\tilde{h} : Form \longrightarrow \{\bot, \top, \Box\bot, \Diamond\top\}$ by definition $\tilde{h}(\phi \wedge \psi) = \tilde{h}(\phi) \wedge \tilde{h}(\psi)$, and $\tilde{h}(\neg\phi) = \neg\tilde{h}(\phi)$, and of course $\tilde{h}(p) = h(p)$ for the variables. $\tilde{h}$ is a homomorphism and again we have $\tilde{h}(\phi) = \Diamond\top \Leftrightarrow M \not\models\!\!> \phi$. (The latter fact can be seen as a consequence of the Boolean algebraic version of $2 \times 2 = 4$.) We also have $\tilde{h}(\gamma) \in \Sigma_1 \Leftrightarrow M \models\!\!> \gamma$. Taking $\tilde{h}$ as an arithmetical interpretation we obtain $\tilde{h}[\Gamma] \subseteq \Sigma_1$ and $\tilde{h}(\phi) \notin \Sigma_1$, i.e. $\Gamma \not\models_{\Sigma_1} \phi$.

C.) $\Vdash_{\Sigma_1} \subseteq \ \vdash_{\Sigma_1}$; Suppose

$$\Gamma \Vdash_{\Sigma_1} \phi. \tag{57}$$

Without loss of generality we can assume that all proposition variables in Γ occur in ϕ. With a model we will mean a truth assignment for these variables. For any model n put $\Gamma_n = \{\gamma \in \Gamma \mid n \not\models \gamma\}$. We will show: $\models \phi \leftrightarrow \bigwedge_{n \not\models \phi} \bigvee \Gamma_n$. So let m be some model.

Assume $m \models \phi$. Suppose for a contradiction that there exists some model n: $n \not\models \phi$ and for all $\gamma \in \Gamma_n$: $m \not\models \gamma$. Then the S-model $\langle \{m, n\} \Vdash \rangle$, where $\Vdash$ simply is $\models$, is a counterexample for (57).

Now assume $m \models \bigwedge_{n \not\models \phi} \bigvee \Gamma_n$. And assume for a contradiction that $m \not\models \phi$. Then $m \models \bigvee \{\gamma \in \Gamma \mid m \not\models \gamma\}$. A contradiction. $\dashv$

6 Conclusion and further research

In this paper we have tried to extend ILM to a logic which has interpolation. A well known counter example for interpolation seemed to emerge because ILM is unable to express Σ_1-ness (see Section 1.2 above and [19]). The main question of this paper is therefore: Is it possible to adjoin to ILM a unary operator Σ_1 and if so does it give us a logic with interpolation.

In preparation for this we formulated a logic ΣL, the language of which contains the usual $\Box$ and the operator Σ_1. The axioms are a trimmed down version of Japaridzes HGL [8]. HGL is a logic which contains, among other things, operators for any class Σ_n, $n \geq 1$. We showed this logic to be modally complete w.r.t. a certain class of Veltman frames (Section 2.3) and arithmetically complete when translating the modal Σ_1 predicate to a formalization of Σ_1-ness and the $\Box$ to a formalization of provability (Section 2.5 and Section 4.2).

Bearing the main goal of this paper in mind it was somewhat of a disappointment to find out that ΣL does not have interpolation (Section 2.4). We carried on nevertheless.

Next we formulated a logic ΣILM. This was (simply) the union of ILM and ΣL where we in addition replaced the (M) schema: $A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C$ by its more natural version $\Sigma_1 C \to (A \rhd B \to A \wedge C \rhd B \wedge C)$. We showed ΣL modally complete w.r.t. Veltman frames very similar to those ILM is shown complete for (in for instance [10]). Basically we showed ΣILM to be complete w.r.t. the class of Veltman frames which is the intersection of those two classes. It was necessary however to fiddle a bit with the forcing relation (definition 3.11). And for our convenience in showing ΣILM arithmetically complete we sharpened the formulation of the modal completeness theorem (Section 3.17).

In my opinion both fiddles are somewhat dirty and one direction for further research could be to try to do without them. This does not give new results directly but might give some more information on the relation between $\rhd$ and Σ_1.

In this context a variation on Veltman frames is noteworthy, namely simplified Visser frames. In simplified Visser frames we consider a binary S instead of a ternary one and the forcing of a formulas $A \rhd B$ is defined as: $w \models A \rhd B \Leftrightarrow \forall v(wRv \text{ and } v \models A \Rightarrow \exists u vSu \text{ and } v \models B)$. The appendix in [19] suggests a very close relation between $\rhd$ and Σ_1 on these models (see comments on the ΣL forcing relation on page 7). We could approach from the other direction and set up a theory for Veltman frames as in [19]. This might give some information on the (possible) necessity of my fiddles.

In order to investigate (the lack of) interpolation for ΣILM we need a notion of bisimulation. We stated in Fact 2.31 that the notion of bisimulation for (the language of) ILM is incomparable to the notion of bisimulation for (the language of) ΣL. Therefore a notion of bisimulation for ΣILM would be quite strong. However in Section 3.6 we managed to show that ΣILM does not have interpolation by using the two separate notions only. The counterexample given was exactly the same as the one that showed ΣL to lack interpolation.

The gap in expressive power of ΣILM (the reason that no interpolant exists) is that of Σ_1-interpolability[7]. The most appealing direction for further research is therefor to investigate the logic of Σ_1-interpolability in combination with ILM.

So the answer to the main question thus reads as follows. Adjoining to ILM a unary operator Σ_1 gives us a relatively simple logic which is modally and arithmetically complete. It (still) lacks interpolation however.

Finally a third direction for further research could be the investigation of fragments of ΣILM. One example of this is given in Section 5.

7 Acknowledgements

Finally I would like to thank a few people who in any way have contributed to this paper.

First my supervisor Dick de Jongh who introduced me to the topic and created the environment in which I could write my paper. Also I like to thank Joost Joosten for various sessions, which were not all directly related to the subject of this paper but nevertheless were very motivating. His proof readings were very useful and speeded up the finalization of this paper considerably. The afternoon with Albert Visser and Dick de Jongh was also very inspiring and gave some insights into rather technical subtleties.

Finally I thank my parents, Marion and Philip Goris, for continuously support. This paper would not have been possible without it.

Symbols[12]

[12] Some symbols occur more than once since they are used for different (but related) purposes.

References

[1] A. Berarducci. The interpretability logic of Peano arithmetic. *Journal of Symbolic Logic*, 55:1059–1089, 1990.

[2] G. Boolos. *The Logic of Provability.* Cambridge University Press, 1993.

[3] C.C. Chang and H.J. Keisler. *Model Theory.* North Holland Publishing Company, 1973.

[4] D.H.J. de Jongh and F. Veltman. Provability logics for relative interpretability. In P. Petkov, editor, *Mathematical logic, Proceedings of the Heyting 1988 Summer School*, pages 31–42. Plenum Press, 1990.

[5] M. de Rijke, P. Blackburn and Y. Venema. *Modal Logic.* Cambridge University Press, 2001.

[6] W. Hodges. *Model Theory.* Cambridge University Press, 1993.

[7] K.N. Ignatiev. The provability logic for Σ_1-interpolability. *Annals of Pure and Applied Logic*, 64:1–25, 1993.

[8] G.K. Japaridze. The logic of the arithmetical hiearchy. *Anals of Pure and Applied Logic*, 66:89–112, 1994.

[9] G.K. Japaridze. A simple proof of arithmetical completeness for Π_1-conservativity logic. *Notre Dame Journal of Formal Logic*, 35:346–354, 1994.

[10] G.K. Japaridze and D. H. J. de Jongh. The logic of provability. In Samuel R. Buss, editor, *Handbook of Proof Theory*, pages 475–536. Elsevier Science Publishers, Amsterdam, 1998.

[11] J.J. Joosten. The propositional Σ_1-logic of PA. Unpublished.

[12] J.J. Joosten. Towards the interpretability logic of all reasonable arithmetical theories. Master's thesis, University of Amsterdam, 1998.

[13] J.J. Joosten and A. Visser. The interpretability logic of *all* reasonable arithmetical theories. *Erkenntnis*, 53(1–2):3–26, 2000.

[14] J.E. Rubin. *Set theory for the mathematician.* Holden day, 1967.

[15] C. Smoryński. *Self-reference and modal logic.* Springer-Verlag, Berlin, 1985.

[16] C. Smoryński. *Logical number theory I.* Springer-Verlag, 1991.

[17] R.M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 25:287–304, 1976.

[18] A. Visser. Interpretability logic. In P. Petkov, editor, *Mathematical logic, Proc. of the Heyting 1988 summer school in Verna, Bulgaria*, pages 175–209, Boston, 1988. Plenum Press.

[19] A. Visser. An overview of interpretability logic. In M. Kracht, M. de Rijke, H. Wansing and M. Zakharyaschev, editors, *Advances in Modal Logic*, volume I, pages 307–359. CSLI Publications, 1998.