Materiae Modales in Logicis Interpretabilitatis

Evan Goris
and
Joost J. Joosten

August 14, 2004

Abstract

In this paper we work out a contruction method for modal models of interpretability logics. Most applications of this construction method involve modal completeness. In the paper, the modal completeness of **IL**, **IL**M, **IL**M_0 and **IL**W^* are proved.

We also expose a modal incompleteness result by showing that **IL**$\mathsf{W}^*\mathsf{P}_0$ is modally incomplete. A new principle R is put foreward. It is shown that R is arithmetically valid in all reasonable arithmetical theories .

Contents

1 Introduction

Interpretability logics are primarily used to describe structural behavior of interpretability between formal mathematical theories. We shall see that the logics come with a good modal semantics that naturally extends the regular modal semantics giving it a dynamical flavor. In this introduction we shall informally describe the project of this paper. Formal definitions are postponed to later sections.

The notion of interpretability that we are primarily interested in, is the notion of relativized interpretability as studied e.g. by Tarski et al in [TMR53]. Roughly, a theory U interprets a theory V –we write $U \rhd V$– if U proves all theorems of V under some structure preserving translation. We allow for relativization of quantifiers. It is defendable to say that U is as least as strong as V if $U \rhd V$. We think that it is clear that interpretations are worth to be studied, as they are omnipresent in both mathematics and meta-mathematics (Langlands Program, relative consistency proofs, undecidability results, Hilberts Programme and so forth).

One approach to the study of interpretability is to study general structural behavior of interpretability. An example of such a structural rule is the transitivity of interpretability. That is, for any U, V and W we have that if $U \rhd V$ and $V \rhd W$, then also $U \rhd W$. As we shall see, modal interpretability logics provide an informative way to support this structural study. Interpretability logics, in a sense, generate all structural rules. Many important questions on interpretability logics have been settled. One of the most prominent open questions at this time is the question of the interpretability logic of all reasonable arithmetical theories . In this paper we make a significant contribution to a solution of this problem. However, a modal characterization still remains an open question.

The main aim of this paper is to establish some modal techniques/toolkit for interpretability logics. Most techniques are aimed at establishing modal completeness results. As we shall see, in the field of interpretability logics, modal completeness can be a sticky business compared to unary modal logics. In this paper we make a first attempt at pulling some (more) thorns out. Significant progress with this respect has also been made by de Jong and Veltman [dJV90].

We have a feeling that the general modal theory of interpretability logics is getting more and more mature. For example, fixed point phenomena and interpolation are quite well understood ([dJV91], [AdJH98], [Vis97]).

Experience tells us that our modal semantics is quite informative and perspicuous. It is even the case that new arithmetical principles can be obtained from modal semantical considerations. An example is our new principle R. We found this principle primarily by modal investigation. Thus, indeed, there is a close match between the modal part and the arithmetical part. It is even possible to embed our modal semantics into some category of models of arithmetic.

Although this paper is mainly a modal investigation, the main questions are still inspired by the arithmetical meaning of our logics. Thus, our investigations will lead to applications concerning arithmetically informative notions like, essentially Σ_1-sentences, self provers and the interpretability logic of all reasonable arithmetical theories .

2 Interpretability logics

2.1 Syntax and conventions

In this paper we shall be mainly interested in interpretability logics, the formulas of which, we write $\mathsf{Form}_{\mathbf{IL}}$, are defined as follows.

$$\mathsf{Form}_{\mathbf{IL}} := \bot \mid \mathsf{Prop} \mid (\mathsf{Form}_{\mathbf{IL}} \to \mathsf{Form}_{\mathbf{IL}}) \mid (\Box\mathsf{Form}_{\mathbf{IL}}) \mid (\mathsf{Form}_{\mathbf{IL}} \rhd \mathsf{Form}_{\mathbf{IL}})$$

Here Prop is a countable set of propositional variables $p, q, r, s, t, p_0, p_1, \ldots$. We employ the usual definitions of the logical operators $\neg, \vee, \wedge$ and $\leftrightarrow$. Also shall we write $\Diamond\varphi$ for $\neg\Box\neg\varphi$. Formulas that start with a $\Box$ are called box-formulas or $\Box$-formulas. Likewise we talk of $\Diamond$-formulas.

From now on we will stay in the realm of interpretability logics. Unless mentioned otherwise, formulas or sentences are formulas of $\mathsf{Form}_{\mathbf{IL}}$. We will write $p \in \varphi$ to indicate that the proposition variable p does occur in φ. A literal is either a propositional variable or the negation of a propositional variable.

In writing formulas we shall omit brackets that are superfluous according to the following reading conventions. We say that the operators $\Diamond$, $\Box$ and $\neg$ bind equally strong. They bind stronger than the equally strong binding $\wedge$ and $\vee$ which in turn bind stronger than $\rhd$. The weakest (weaker than $\rhd$) binding connectives are $\to$ and $\leftrightarrow$. We shall also omit outer brackets. Thus, we shall write $A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C$ instead of $((A \rhd B) \to ((A \wedge (\Box C)) \rhd (B \wedge (\Box C))))$.

A schema of interpretability logic is syntactically like a formula. They are used to generate formulae that have a specific form. We will not be specific about the syntax of schemata as this is similar to that of formulas. Below, one can think of A, B and C as place holders.

The rule of Modus Ponens allows one to conclude B from premises $A \to B$ and A. The rule of Necessitation allows one to conclude $\Box A$ from the premise A.

Definition 2.1. The logic **IL** is the smallest set of formulas being closed under the rules of Necessitation and of Modus Ponens, that contains all tautological formulas and all instantiations of the following axiom schemata.

L1 $\Box(A \to B) \to (\Box A \to \Box B)$

L2 $\Box A \to \Box\Box A$

L3 $\Box(\Box A \to A) \to \Box A$

J1 $\Box(A \to B) \to A \rhd B$

J2 $(A \rhd B) \wedge (B \rhd C) \to A \rhd C$

J3 $(A \rhd C) \wedge (B \rhd C) \to A \vee B \rhd C$

J4 $A \rhd B \to (\Diamond A \to \Diamond B)$

J5 $\Diamond A \rhd A$

We will write $\mathbf{IL} \vdash \varphi$ for $\varphi \in \mathbf{IL}$. An **IL**-derivation or **IL**-proof of φ is a finite sequence of formulae ending on φ, each being a logical tautology, an instantiation of one of the axiom schemata of **IL**, or the result of applying either Modus Ponens or Necessitation to formulas earlier in the sequence. Clearly, $\mathbf{IL} \vdash \varphi$ iff. there is an **IL**-proof of φ.

Sometimes we will write $\mathbf{IL} \vdash \varphi \to \psi \to \chi$ as short for $\mathbf{IL} \vdash \varphi \to \psi$ & $\mathbf{IL} \vdash \psi \to \chi$. Similarly for $\rhd$. We adhere to a similar convention when we employ binary relations. Thus, $xRyS_xz \Vdash B$ is short for xRy & yS_xz & $z \Vdash B$, and so on.

Sometimes we will consider the part of **IL** that does not contain the $\rhd$-modality. This is the well-known provability logic **GL**, whose axiom schemata are L1-L3. The axiom schema L3 is often referred to as Löb's axiom.

Lemma 2.2.

1. $\mathbf{IL} \vdash \Box A \leftrightarrow \neg A \rhd \bot$
2. $\mathbf{IL} \vdash A \rhd A \wedge \Box\neg A$
3. $\mathbf{IL} \vdash A \vee \Diamond A \rhd A$

Proof. All of these statements have very easy proofs. We give an informal proof of the second statement. Reason in **IL**. It is easy to see $A \rhd (A \wedge \Box\neg A) \vee (A \wedge \Diamond A)$. By L3 we get $\Diamond A \to \Diamond(A \wedge \Box\neg A)$. Thus, $A \wedge \Diamond A \rhd \Diamond(A \wedge \Box\neg A)$ and by J5 we get $\Diamond(A \wedge \Box\neg A) \rhd A \wedge \Box\neg A$. As certainly $A \wedge \Box\neg A \rhd A \wedge \Box\neg A$ we have that $(A \wedge \Box\neg A) \vee (A \wedge \Diamond A) \rhd A \wedge \Box\neg A$ and the result follows from transitivity of $\rhd$. ⊣

Apart from the axiom schemata exposed in Definition 2.1 we will on occassion consider other axiom schemata too.

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

P $A \rhd B \to \Box(A \rhd B)$

M_0 $A \rhd B \to \Diamond A \wedge \Box C \rhd B \wedge \Box C$

W $A \rhd B \to A \rhd B \wedge \Box\neg A$

W^* $A \rhd B \to B \wedge \Box C \rhd B \wedge \Box C \wedge \Box\neg A$

P_0 $A \rhd \Diamond B \to \Box(A \rhd B)$

R $A \rhd B \to \neg(A \rhd \neg C) \rhd B \wedge \Box C$

If X is a set of axiom schemata we will denote by $\mathbf{IL}\mathsf{X}$ the logic that arises by adding the axiom schemata in X to $\mathbf{IL}$. Thus, $\mathbf{IL}\mathsf{X}$ is the smallest set of formulas being closed under the rules of Modus Ponens and Necessitation and containing all tautologies and all instantiations of the axiom schemata of $\mathbf{IL}$ ($\mathsf{L1}$-$\mathsf{J5}$) and of the axiom schemata of X. Instead of writing $\mathbf{IL}\{\mathsf{M_0}, \mathsf{W}\}$ we will write $\mathbf{IL}\mathsf{M_0W}$ and so on.

We write $\mathbf{IL}\mathsf{X} \vdash \varphi$ for $\varphi \in \mathbf{IL}\mathsf{X}$. An $\mathbf{IL}\mathsf{X}$-derivation or $\mathbf{IL}\mathsf{X}$-proof of φ is a finite sequence of formulae ending on φ, each being a logical tautology, an instantiation of one of the axiom schemata of $\mathbf{IL}\mathsf{X}$, or the result of applying either Modus Ponens or Necessitation to formulas earlier in the sequence. Again, $\mathbf{IL}\mathsf{X} \vdash \varphi$ iff. there is an $\mathbf{IL}\mathsf{X}$-proof of φ. For a schema Y, we write $\mathbf{IL}\mathsf{X} \vdash \mathsf{Y}$ if $\mathbf{IL}\mathsf{X}$ proves every instantiation of Y.

Definition 2.3. Let Γ be a set of formulas. We say that φ is provable from Γ in $\mathbf{IL}\mathsf{X}$ and write $\Gamma \vdash_{\mathbf{IL}\mathsf{X}} \varphi$, iff. there is a finite sequence of formulae ending on φ, each being a theorem of $\mathbf{IL}\mathsf{X}$, a formula from Γ, or the result of applying Modus Ponens to formulas earlier in the sequence.

Clearly we have $\varnothing \vdash_{\mathbf{IL}\mathsf{X}} \varphi \Leftrightarrow \mathbf{IL}\mathsf{X} \vdash \varphi$. In the sequel we will often write just $\Gamma \vdash \varphi$ instead of $\Gamma \vdash_{\mathbf{IL}\mathsf{X}} \varphi$ if the context allows us so. It is well known that we have a deduction theorem for this notion of derivability.

Lemma 2.4 (Deduction theorem). $\Gamma, A \vdash_{\mathbf{IL}\mathsf{X}} B \Leftrightarrow \Gamma \vdash_{\mathbf{IL}\mathsf{X}} A \to B$

Proof. "$\Leftarrow$" is obvious and "$\Rightarrow$" goes by induction on the length n of the $\mathbf{IL}\mathsf{X}$-proof σ of B from Γ, A.

If $n{>}1$, then $\sigma = \tau, B$, where B is obtained from some C and $C \to B$ occurring earlier in τ. Thus we can find subsequences τ' and τ'' of τ such that τ', C and $\tau'', C \to B$ are $\mathbf{IL}\mathsf{X}$-proofs from Γ, A. By the induction hypothesis we find $\mathbf{IL}\mathsf{X}$-proofs from Γ of the form $\sigma', A \to C$ and $\sigma'', A \to (C \to B)$. We now use the tautology $(A \to (C \to B)) \to ((A \to C) \to (A \to B))$ to get an $\mathbf{IL}\mathsf{X}$-proof of $A \to B$ from Γ. $\dashv$

Definition 2.5. A set Γ is $\mathbf{IL}\mathsf{X}$-consistent iff. $\Gamma \nvdash_{\mathbf{IL}\mathsf{X}} \bot$. An $\mathbf{IL}\mathsf{X}$-consistent set is maximal $\mathbf{IL}\mathsf{X}$-consistent if for any φ, either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$.

Lemma 2.6. *Every* $\mathbf{IL}\mathsf{X}$*-consistent set can be extended to a maximal* $\mathbf{IL}\mathsf{X}$*-consistent one.*

Proof. This is Lindebaums lemma for $\mathbf{IL}\mathsf{X}$. We can just do the regular argument as we have the deduction theorem. Note that there are countably many different formulas. $\dashv$

We will often abbreviate "maximal consistent set" by MCS and refrain from explicitly mentioning the logic $\mathbf{IL}\mathsf{X}$ when the context allows us to do so. We define three useful relations on MCS's, the *successor* relation $\prec$, the C*-critical successor* relation $\prec_C$ and the *Box-inclusion* relation $\subseteq_\Box$.

Definition 2.7. Let Γ and Δ denote maximal $\mathbf{IL}\mathsf{X}$-consistent sets.

- $\Gamma \prec \Delta := \Box A \in \Gamma \Rightarrow A, \Box A \in \Delta$
- $\Gamma \prec_C \Delta := A \rhd C \in \Gamma \Rightarrow \neg A, \Box\neg A \in \Delta$
- $\Gamma \subseteq_\Box \Delta := \Box A \in \Gamma \Rightarrow \Box A \in \Delta$

It is clear that $\Gamma \prec_C \Delta \Rightarrow \Gamma \prec \Delta$. For, if $\Box A \in \Gamma$ then $\neg A \rhd \bot \in \Gamma$. Also $\bot \rhd C \in \Gamma$, whence $\neg A \rhd C \in \Gamma$. If now $\Gamma \prec_C \Delta$ then $A, \Box A \in \Delta$, whence $\Gamma \prec \Delta$. It is also clear that $\Gamma \prec_C \Delta \prec \Delta' \Rightarrow \Gamma \prec_C \Delta'$.

Lemma 2.8. *Let Γ and Δ denote maximal* **IL**X*-consistent sets. We have $\Gamma \prec \Delta$ iff. $\Gamma \prec_\bot \Delta$.*

Proof. Above we have seen that $\Gamma \prec_A \Delta \Rightarrow \Gamma \prec \Delta$. For the other direction suppose now that $\Gamma \prec \Delta$. If $A \rhd \bot \in \Gamma$ then, by Lemma 2.2.1, $\Box\neg A \in \Gamma$ whence $\neg A, \Box\neg A \in \Delta$. $\dashv$

2.2 Semantics

Interpretability logics come with a Kripke-like semantics. As the signature of our language is countable, we shall only consider countable models.

Definition 2.9. An **IL**-frame is a triple $\langle W, R, S\rangle$. Here W is a non-empty countable universe, R is a binary relation on W and S is a set of binary relations on W, indexed by elements of W. The R and S satisfy the following requirements.

1. R is conversely well-founded[1]
2. xRy & $yRz \rightarrow xRz$
3. $yS_xz \rightarrow xRy$ & xRz
4. $xRy \rightarrow yS_xy$
5. $xRyRz \rightarrow yS_xz$
6. $uS_xvS_xw \rightarrow uS_xw$

IL-frames are sometimes also called Veltman frames. We will on occasion speak of R or S_x transitions instead of relations. If we write ySz, we shall mean that yS_xz for some x. W is sometimes called the universe, or domain, of the frame and its elements are referred to as worlds or nodes. With $x{\upharpoonright}$ we shall denote the set $\{y \in W \mid xRy\}$. We will often represent S by a ternary relation in the canonical way, writing $\langle x, y, z\rangle$ for yS_xz.

Definition 2.10. An **IL**-model is a quadruple $\langle W, R, S, \Vdash\rangle$. Here $\langle W, R, S, \rangle$ is an **IL**-frame and $\Vdash$ is a subset of $W \times \mathsf{Prop}$. We write $w \Vdash p$ for $\langle w, p\rangle \in \Vdash$. As usual, $\Vdash$ is extended to a subset $\widetilde{\Vdash}$ of $W \times \mathsf{Form}_{\mathbf{IL}}$ by demanding the following.

[1] A relation R on W is called conversely well-founded if every non-empty subset of W has an R-maximal element.

- $w \widetilde{\Vdash} p$ iff. $w \Vdash p$ for $p \in \mathsf{Prop}$
- $w \not\widetilde{\Vdash} \bot$
- $w \widetilde{\Vdash} A \to B$ iff. $w \not\widetilde{\Vdash} A$ or $w \widetilde{\Vdash} B$
- $w \widetilde{\Vdash} \Box A$ iff. $\forall v\ (wRv \Rightarrow v \widetilde{\Vdash} A)$
- $w \widetilde{\Vdash} A \rhd B$ iff. $\forall u\ (wRu \wedge u \widetilde{\Vdash} A \Rightarrow \exists v\ (uS_w v \widetilde{\Vdash} B))$

Note that $\widetilde{\Vdash}$ is completely determined by $\Vdash$. Thus we will denote $\widetilde{\Vdash}$ also by $\Vdash$. We call $\Vdash$ a forcing relation. The $\Vdash$-relation depends on the model M. If necessary, we will write $M, w \Vdash \varphi$, if not, we will just write $w \Vdash \varphi$. In this case we say that φ holds at w, or that φ is forced at w. We say that *p is in the range of* $\Vdash$ if $w \Vdash p$ for some w.

If $F = \langle W, R, S \rangle$ is an **IL**-frame, we will write $x \in F$ to denote $x \in W$ and similarly for **IL**-models. Attributes on F will be inherited by its constituent parts. For example $F_i = \langle W_i, R_i, S_i \rangle$. Often however we will write $F_i \models xRy$ instead of $F_i \models xR_iy$ and likewise for the S-relation. This notation is consistent with notation in first order logic where the symbol R is interpreted in the structure F_i as R_i.

If $M = \langle W, R, S, \Vdash \rangle$, we say that M is based on the frame $\langle W, R, S \rangle$ and we call $\langle W, R, S \rangle$ its underlying frame.

If Γ is a set of formulas, we will write $M, x \Vdash \Gamma$ as short for $\forall \gamma{\in}\Gamma\ M, x \Vdash \gamma$. We have similar reading conventions for frames and for validity.

Definition 2.11 (Generated Submodel). Let $M = \langle W, R, S, \Vdash \rangle$ be an **IL**-model and let $m \in M$. We define $m{\upharpoonright}{*}$ to be the set $\{x \in W \mid x{=}m \vee mRx\}$. By $M{\upharpoonright}m$ we denote the submodel generated by m defined as follows.

$$M{\upharpoonright}m := \langle m{\upharpoonright}{*}, R \cap (m{\upharpoonright}{*})^2, \bigcup_{x \in m\upharpoonright *} S_x \cap (m{\upharpoonright}{*})^2, \Vdash \cap (m{\upharpoonright}{*} \times \mathsf{Prop}) \rangle$$

Lemma 2.12 (Generated Submodel Lemma). *Let M be an* **IL***-model and let $m \in M$. For all formulas φ and all $x \in m{\upharpoonright}{*}$ we have that*

$$M{\upharpoonright}m, x \Vdash \varphi \quad \textit{iff.} \quad M, x \Vdash \varphi.$$

Proof. By an easy induction on the complexity of φ. ⊣

We say that an **IL**-model makes a formula φ true, and write $M \models \varphi$, if φ is forced in all the nodes of M. In a formula we write

$$M \models \varphi :\Leftrightarrow \forall\, w{\in}M\ w \Vdash \varphi.$$

If $F = \langle W, R, S \rangle$ is an **IL**-frame and $\Vdash$ a subset of $W \times \mathsf{Prop}$, we denote by $\langle W, \Vdash \rangle$ the **IL**-model that is based on F and has forcing relation $\Vdash$. We say that a frame F makes a formula φ true, and write $F \models \varphi$, if any model based on F makes φ true. In a second-order formula:

$$F \models \varphi :\Leftrightarrow \forall \Vdash \ \langle F, \Vdash \rangle \models \varphi$$

We say that an **IL**-model or frame makes a scheme true if it makes all its instantiations true. If we want to express this by a formula we should have a means to quantify over all instantiations. For example, we could regard an instantiation of a scheme X as a substitution σ carried out on X resulting in X^σ. We do not wish to be very precise here, as it is clear what is meant. Our definitions thus read

$$F \models \mathsf{X} \text{ iff. } \forall \sigma \ F \models \mathsf{X}^\sigma$$

for frames F, and

$$M \models \mathsf{X} \text{ iff. } \forall \sigma \ M \models \mathsf{X}^\sigma$$

for models M. Sometimes we will also write $F \models \mathbf{IL}\mathsf{X}$ for $F \models \mathsf{X}$.

It turns out that checking the validity of a scheme on a frame is fairly easy. If X is some scheme[2], let τ be some base substitution that sends different placeholders to different propositional variables.

Lemma 2.13. *Let X be a scheme, and τ be a corresponding base substitution as described above. Let F be an* **IL***-frame. We have*

$$F \models \mathsf{X}^\tau \Leftrightarrow \forall \sigma \ F \models \mathsf{X}^\sigma.$$

Proof. If $\forall \sigma \ F \models \mathsf{X}^\sigma$, then certainly $F \models \mathsf{X}^\tau$, thus we should concentrate on the other direction. Thus, assuming $F \models \mathsf{X}^\tau$ we fix some σ and $\Vdash$ and set out to prove $\langle F, \Vdash \rangle \models \mathsf{X}^\sigma$. We define another forcing relation $\Vdash'$ on F by saying that for any place holder A in X we have

$$w \Vdash' \tau(A) :\Leftrightarrow \langle F, \Vdash \rangle \models \sigma(A)$$

By induction on the complexity of a subscheme[3] Y of X we can now prove

$$\langle F, \Vdash' \rangle, w \Vdash' \mathsf{Y}^\tau \Leftrightarrow \langle F, \Vdash \rangle, w \Vdash \mathsf{Y}^\sigma.$$

By our assumption we get that $\langle F, \Vdash \rangle, w \Vdash \mathsf{X}^\sigma$. ⊣

If χ is some formula in first, or higher, order predicate logic, we will evaluate $F \models \chi$ in the standard way. In this case F is considered as a structure of first or higher order predicate logic. We will not be too formal about these matters as the context will always dict us which reading to choose.

[2]Or a set of schemata. All of our reasoning generalizes without problems to sets of schemata. We will therefore no longer mention the distinction.

[3]It is clear what this notion should be.

Definition 2.14. Let X be a scheme of interpretability logic. We say that a formula $\mathcal{C}$ in first or higher order predicate logic is a frame condition of X if

$$F \models \mathcal{C} \quad \text{iff.} \ F \models \mathsf{X}.$$

The $\mathcal{C}$ in Definition 2.14 is also called the frame condition of the logic $\mathbf{IL}\mathsf{X}$. A frame satisfying the $\mathbf{IL}\mathsf{X}$ frame condition is often called an $\mathbf{IL}\mathsf{X}$-frame. In case no such frame condition exists, an $\mathbf{IL}\mathsf{X}$-frame resp. model is just a frame resp. model, validating X.

The semantics for interpretability logics is good in the sense that we have the necessary soundness results.

Lemma 2.15 (Soundness). $\mathbf{IL} \vdash \varphi \Rightarrow \forall F \ F \models \varphi$

Proof. By induction on the length of an $\mathbf{IL}$-proof of φ. The requirements on R and S in Definition 2.9 are precisely such that the axiom schemata hold. Note that all axiom schemata have their semantical counterpart except for the schema $(A \rhd C) \wedge (B \rhd C) \rightarrow A \vee B \rhd C$. ⊣

Lemma 2.16 (Soundness). *Let $\mathcal{C}$ be the frame condition of the logic $\mathbf{IL}\mathsf{X}$. We have that*

$$\mathbf{IL}\mathsf{X} \vdash \varphi \Rightarrow \forall F \ (F \models \mathcal{C} \Rightarrow F \models \varphi).$$

Proof. As that of Lemma 2.15, plugging in the definition of the frame condition at the right places. Note that we only need the direction $F \models \mathcal{C} \Rightarrow F \models X$ in the proof. ⊣

Corollary 2.17. *Let M be a model satisfying the $\mathbf{IL}\mathsf{X}$ frame condition, and let $m \in M$. We have that $\Gamma := \{\varphi \mid M, m \Vdash \varphi\}$ is a maximal $\mathbf{IL}\mathsf{X}$-consistent set.*

Proof. Clearly $\perp \notin \Gamma$. Also $A \in \Gamma$ or $\neg A \in \Gamma$. By the soundness lemma, Lemma 2.16, we see that Γ is closed under $\mathbf{IL}\mathsf{X}$ consequences. ⊣

Lemma 2.18. *Let M be a model such that $\forall\, w{\in}M \ \ w \Vdash \mathbf{IL}\mathsf{X}$ then $\mathbf{IL}\mathsf{X} \vdash \varphi \Rightarrow M \models \varphi$.*

Proof. By induction on the derivation of φ. ⊣

A modal logic $\mathbf{IL}\mathsf{X}$ with frame condition $\mathcal{C}$ is called complete if we have the implication the other way round too. That is,

$$\forall F \ (F \models \mathcal{C} \Rightarrow F \models \varphi) \Rightarrow \mathbf{IL}\mathsf{X} \vdash \varphi.$$

A major concern of this paper is the question whether a given modal logic $\mathbf{IL}\mathsf{X}$ is complete.

Definition 2.19. $\Gamma \Vdash_{\mathbf{IL}\mathsf{X}} \varphi$ iff. $\forall M \ M \models \mathbf{IL}\mathsf{X} \Rightarrow (\forall\, m{\in}M \ [M, m \Vdash \Gamma \Rightarrow M, m \Vdash \varphi])$

Lemma 2.20. *Let Γ be a finite set of formulas and let* $\mathbf{IL}\mathsf{X}$ *be a complete logic. We have that* $\Gamma \vdash_{\mathbf{IL}\mathsf{X}} \varphi$ *iff.* $\Gamma \Vdash_{\mathbf{IL}\mathsf{X}} \varphi$.

Proof. Trivial. By the deduction theorem $\Gamma \vdash_{\mathbf{IL}\mathsf{X}} \varphi \Leftrightarrow \vdash_{\mathbf{IL}\mathsf{X}} \bigwedge \Gamma \to \varphi$. By our assumption on completeness we get the result. Note that the requirement that Γ be finite is necessary, as our modal logics are in general not compact (see also Section 3.1). $\dashv$

Often we shall need to compare different frames or models. If $F = \langle W, R, S\rangle$ and $F' = \langle W', R', S'\rangle$ are frames, we say that F is a subframe of F' and write $F \subseteq F'$, if $W \subseteq W'$, $R \subseteq R'$ and $S \subseteq S'$. Here $S \subseteq S'$ is short for $\forall\, w{\in}W\ (S_w \subseteq S'_w)$.

2.3 Arithmetic

As with (almost) all interesting occurrences of modal logic, interpretability logics are used to study a hard mathematical notion. Interpretability logics, as their name slightly suggests, are used to study the notion of formal interpretability. In this subsection we shall very briefly say what this notion is and how modal logic is used to study it.

We are interested in first order theories in the language of arithmetic. All theories we will consider will thus be arithmetical theories. Moreover, we want our theories to have a certain minimal strength. That is, they should contain a certain core theory, say $\mathsf{I}\Delta_0{+}\Omega_1$ from [HP93]. This will allow us to do reasonable coding of syntax. We call these theories reasonable arithmetical theories.

Once we can code syntax, we can write down a decidable predicate $\mathsf{Proof}_T(p, \varphi)$ that holds on the standard model precisely when p is a T-proof of φ.[4] We get a provability predicate by quantifying existentially, that is, $\mathsf{Prov}_T(\varphi) := \exists p\ \mathsf{Proof}_T(p, \varphi)$.

We can use these coding techniques to code the notion of formal interpretability too. Roughly, a theory U interprets a theory V if there is some sort of translation so that every theorem of V is under that translation also a theorem of U.

Definition 2.21. Let U and V be reasonable arithmetical theories. An interpretation j from V in U is a pair $\langle \delta, F\rangle$. Here, δ is called a domain specifier. It is a formula with one free variable. The F is a map that sends an n-ary relation symbol of V to a formula of U with n free variables. (We treat functions and constants as relations with additional properties.) The interpretation j induces a translation from formulas φ of V to formulas φ^j of U by replacing relation symbols by their corresponding formulas and by relativizing quantifiers to δ. We have the following requirements.

- $(R(\vec{x}))^j = F(R)(\vec{x})$

[4] We take the liberty to not make a distinction between a syntactical object and its code.

- The translation induced by j commutes with the boolean connectives. Thus, for example, $(\varphi \vee \psi)^j = \varphi^j \vee \psi^j$. In particular $(\bot)^j = (\bigvee_{\varnothing})^j = \bigvee_{\varnothing} = \bot$
- $(\forall x\ \varphi)^j = \forall x\ (\delta(x) \to \varphi^j)$
- $V \vdash \varphi \Rightarrow U \vdash \varphi^j$

We say that V is interpretable in U if there exists an interpretation j of V in U.

Using the $\mathsf{Prov}_T(\varphi)$ predicate, it is possible to code the notion of formal interpretability in arithmetical theories. This gives rise to a formula $\mathsf{Int}_T(\varphi, \psi)$, to hold on the standard model precisely when $T + \psi$ is interpretable in $T + \varphi$. This formula is related to the modal part by means of arithmetical realizations.

Definition 2.22. An arithmetical realization $*$ is a mapping that assigns to each propositional variable an arithmetical sentence. This mapping is extended to all modal formulas in the following way.

- $(\varphi \vee \psi)^* = \varphi^* \vee \psi^*$ and likewise for other boolean connectives. In particular $\bot^* = (\bigvee_{\varnothing})^* = \bigvee_{\varnothing} = \bot$.
- $(\Box\varphi)^* = \mathsf{Prov}_T(\varphi^*)$
- $(\varphi \rhd \psi)^* = Int_T(\varphi^*, \psi^*)$

From now on, the $*$ will always range over realizations. Often we will write $\Box_T\varphi$ instead of $\mathsf{Prov}_T(\varphi)$ or just even $\Box\varphi$. The $\Box$ can thus denote both a modal symbol and an arithmetical formula. For the $\rhd$-modality we adopt a similar convention. We are confident that no confusion will arise from this.

Definition 2.23. An interpretability principle of a theory T is a modal formula φ that is provable in T under any realization. That is, $\forall *\ T \vdash \varphi^*$. The interpretability logic of a theory T, we write $\mathbf{IL}(\mathrm{T})$, is the set of all interpretability principles.

Likewise, we can talk of the set of all provability principles of a theory T, denoted by $\mathbf{PL}(\mathrm{T})$. Since the famous result by Solovay, $\mathbf{PL}(\mathrm{T})$ is known for a large class of theories T.

Theorem 2.24 (Solovay [Sol76]). $\mathbf{PL}(\mathrm{T}) = \mathbf{GL}$ *for any reasonable arithmetical theory* T.

For two classes of theories, $\mathbf{IL}(\mathrm{T})$ is known.

Definition 2.25. A theory T is reflexive if it proves the consistency of any of its finite subtheories. It is essentially reflexive if any finite extension of it is reflexive.

Theorem 2.26 (Berarducci [Ber90], Shavrukov [Sha88]). *If* T *is an essentially reflexive theory, then* $\mathbf{IL}(\mathrm{T}) = \mathbf{IL}\mathsf{M}$.

Theorem 2.27 (Visser [Vis90]). *If T is finitely axiomatizable, then* $\mathbf{IL}(\mathrm{T}) = \mathbf{IL}\mathsf{P}$.

Definition 2.28. The interpretability logic of all reasonable arithmetical theories, we write $\mathbf{IL}(\mathrm{All})$, is the set of formulas φ such that $\forall T \, \forall * \;\; T \vdash \varphi^*$. Here the T ranges over all the reasonable arithmetical theories.

For sure $\mathbf{IL}(\mathrm{All})$ should be in the intersection of $\mathbf{IL}\mathsf{M}$ and $\mathbf{IL}\mathsf{P}$. Up to now, $\mathbf{IL}(\mathrm{All})$ is unknown. In [JV00] it is conjectured to be $\mathbf{IL}\mathsf{P_0W^*}$. It is one of the major open problems in the field of interpretability logics, to characterize $\mathbf{IL}(\mathrm{All})$ in a modal way.

We conclude this subsection with a definition of the arithmetical hierarchy. This definition is needed in Section 7.

Definition 2.29. Inductively the following classes of arithmetical formulae are defined.

- Arithmetical formulas with only bounded quantifiers in it are called Δ_0, Σ_0 or Π_0-formulas.
- If φ is a Π_n or Σ_{n+1}-formula, then $\exists x \; \varphi$ is a Σ_{n+1}-formula.
- If φ is a Σ_n or Π_{n+1}-formula, then $\forall x \; \varphi$ is a Π_{n+1}-formula.

Definition 2.30. Let φ be an arithmetical formula.

- $\varphi \in \Pi_n(T)$ iff. $\exists \, \pi{\in}\Pi_n \; T \vdash \varphi \leftrightarrow \pi$
- $\varphi \in \Sigma_n(T)$ iff. $\exists \, \sigma{\in}\Sigma_n \; T \vdash \varphi \leftrightarrow \sigma$
- $\varphi \in \Delta_n(T)$ iff. $\exists \, \pi{\in}\Pi_n \;\&\; \exists \, \sigma{\in}\Sigma_n \; T \vdash (\varphi \leftrightarrow \pi) \wedge (\varphi \leftrightarrow \sigma)$

Sometimes, if no confusion can arise, we will write Σ_n!-formulas instead of Σ_n-formulas and Σ_n-formulas instead of $\Sigma_n(T)$-formulas.

3 General exposition of the construction method

Most of the applications of the construction method deal with modal completeness of a certain logic $\mathbf{IL}\mathsf{X}$. More precisely, showing that a logic $\mathbf{IL}\mathsf{X}$ is modally complete amounts to constructing, or finding, whenever $\mathbf{IL}\mathsf{X} \nvdash \varphi$, a model M and an $x \in M$ such that $M, x \Vdash \neg\varphi$. We will employ our construction method for this particular model construction.

In this section, we will not always give precise definitions of the notions we work with. All the definitions can be found in Section 4.

3.1 The main ingredients of the construction method

As we mentioned above, a modal completeness proof of a logic $\mathbf{IL}\mathsf{X}$ amounts to a uniform model construction to obtain $M, x \Vdash \neg\varphi$ for $\mathbf{IL}\mathsf{X} \nvdash \varphi$. If $\mathbf{IL}\mathsf{X} \nvdash \varphi$, then $\{\neg\varphi\}$ is an $\mathbf{IL}\mathsf{X}$-consistent set and thus, by a version of Lindenbaum's Lemma (Lemma 2.6), it is extendible to a maximal $\mathbf{IL}\mathsf{X}$-consistent set. On the other hand, once we have an $\mathbf{IL}\mathsf{X}$-model $M, x \Vdash \neg\varphi$, we can find, by Corollary 2.17 a maximal $\mathbf{IL}\mathsf{X}$-consistent set Γ with $\neg\varphi \in \Gamma$. This Γ can simply be defined as the set of all formulas that hold at x.

To go from a maximal $\mathbf{IL}\mathsf{X}$-consistent set to a model is always the hard part. This part is carried out in our construction method. In this method, the maximal consistent set is somehow partly unfolded to a model.

Often in these sort of model constructions, the worlds in the model are MCS's. For propositional variables one then defines $x \Vdash p$ iff. $p \in x$. In the setting of interpretability logics it is sometimes inevitable to use the same MCS in different places in the model.[5] Therefore we find it convenient not to identify a world x with a MCS, but rather label it with a MCS $\nu(x)$. However, we will still write sometimes $\varphi \in x$ instead of $\varphi \in \nu(x)$.

One complication in unfolding a MCS to a model lies in the incompactness of the modal logics we consider. This, in turn, is due to the fact that some frame conditions are not expressible in first order logic. As an example we can consider the following set.[6]

$$\Gamma := \{\Diamond p_0\} \cup \{\Box(p_i \to \Diamond p_{i+1}) \mid i \in \omega\}$$

Clearly, Γ is a $\mathbf{GL}$-consistent set, and any finite part of it is satisfiable in some world in some model. However, it is not hard to see that in no $\mathbf{IL}$-model all of Γ can hold simultaneously in some world in it.

If M is an $\mathbf{IL}\mathsf{X}$-model and $x \in M$, then $\{\varphi \mid M, x \Vdash \varphi\}$ is a MCS. By definition (and abuse of notation) we see that

$$\forall x\, [x \Vdash \varphi \quad \text{iff.} \quad \varphi \in x].$$

We call this equivalence a truth lemma. (See for example Definition 4.5 for a more precise formulation.) In all completeness proofs a model is defined or constructed in which some form of a truth lemma holds. Now, by the observed incompactness phenomenon, we can not expect that for every MCS, say Γ, we can find a model "containing" Γ for which a truth lemma holds in full generality. There are various ways to circumvent this complication. Often one considers truncated parts of maximal consistent sets which are finite. In choosing how to truncate, one is driven by two opposite forces.

[5] As the truth definition of $A \rhd B$ has a $\forall\exists$ character, the corresponding notion of bisimulation is rather involved. As a consequence there is in general no obvious notion of a minimal bisimular model, contrary to the case of provability logics. This causes the necessity of several occurrences of MCS's.

[6] This example comes from Fine and Rautenberg and is treated in Chapter 7 of [Boo93].

On the one hand this truncated part should be small. It should be at least finite so that the incompactness phenomenon is blocked. The finiteness is also a desideratum if one is interested in the decidability of a logic.

On the other hand, the truncated part should be large. It should be large enough to admit inductive reasoning to prove a truth lemma. For this, often closure under subformulas and single negation suffices. Also, the truncated part should be large enough so that MCS's contain enough information to do the required calculation. For this, being closed under subformulas and single negations does not, in general, suffice. Examples of these sort of calculation are Lemma 6.7 and Lemma 8.16.

In our approach we take the best of both opposites. That is, we do not truncate at all. Like this, calculation becomes uniform, smooth and relatively easy. However, we demand a truth lemma to hold only for finitely many formulas.

The question is now, how to unfold the MCS containing $\neg\varphi$ to a model where $\neg\varphi$ holds in some world. We would have such a model if a truth lemma holds w.r.t. a finite set $\mathcal{D}$ containing $\neg\varphi$.

Proving that a truth lemma holds is usually done by induction on the complexity of formulas. As such, this is a typical "bottom up" or "inside out" activity. On the other hand, unfolding, or reading off, the truth value of a formula is a typical "top down" or "outside in" activity.

Yet, we do want to gradually build up a model so that we get closer and closer to a truth lemma. But, how could we possibly measure that we come closer to a truth lemma? Either everything is in place and a truth lemma holds, or a truth lemma does not hold, in which case it seems unclear how to measure to what extend it does not hold.

The gradually building up a model will take place by consecutively adding bits and pieces to the MCS we started out with. Thus somehow, we do want to measure that we come closer to a truth lemma by doing so. Therefore, we switch to an alternative forcing relation $\Vdash\!\sim$ that follows the "outside in" direction that is so characteristic to the evaluation of $x \Vdash \varphi$, but at the same time incorporates the necessary elements of a truth lemma.

$x \Vdash\!\sim p$	iff.	$p \in x$	for propositional variables p
$x \Vdash\!\sim \varphi \wedge \psi$	iff.	$x \Vdash\!\sim \varphi$ & $x \Vdash\!\sim \psi$	and likewise for other boolean connectives
$x \Vdash\!\sim \varphi \rhd \psi$	iff.	$\forall y\ [xRy \wedge \varphi \in x \rightarrow \exists z\ (yS_x z \wedge \psi \in z)]$	

If $\mathcal{D}$ is a set of sentences that is closed under subformulas and single negations, then it is not hard to see that (see Lemma 4.9)

$$\forall x \forall\, \varphi{\in}\mathcal{D}\ [x \Vdash\!\sim \varphi \text{ iff. } \varphi \in x] \quad (*)$$

is equivalent to

$$\forall x \forall\, \varphi{\in}\mathcal{D}\ [x \Vdash \varphi \text{ iff. } \varphi \in x]. \quad (**)$$

Thus, if we want to obtain a truth lemma for a finite set $\mathcal{D}$ that is closed under single negations and subformulas, we are done if we can obtain $(*)$. But now it is clear how we can at each step measure that we come closer to a truth lemma. This brings us to the definition of problems and deficiencies.

A problem is some formula $\neg(\varphi \rhd \psi) \in x \cap \mathcal{D}$ such that $x \Vdash\!\!\!\!/ \neg(\varphi \rhd \psi)$. We define a deficiency to be a configuration such that $\varphi \rhd \psi \in x \cap \mathcal{D}$ but $x \Vdash\!\!\!\!/ \varphi \rhd \psi$. It now becomes clear how we can successively eliminate problems and deficiencies.

A deficiency $\varphi \rhd \psi \in x \cap \mathcal{D}$ is a deficiency because there is some y (or maybe more of them) with xRy, and $\varphi \in y$, but for no z with yS_xz, we have $\psi \in z$. This can simply be eliminated by adding a z with yS_xz and $\psi \in z$.

A problem $\neg(\varphi \rhd \psi) \in x \cap \mathcal{D}$ can be eliminated by adding a completely isolated y to the model with xRy and $\varphi, \neg\psi \in y$. As y is completely isolated, $yS_xz \Rightarrow z = y$ and thus indeed, it is not possible to reach a world where ψ holds. Now here is one complication.

We want that a problem or a deficiency, once eliminated, can never re-occur. For deficiencies this complication is not so severe, as the quantifier complexity is $\forall\exists$. Thus, any time "a deficiency becomes active", we can immediately deal with it.

With the elimination of a problem, things are more subtle. When we introduced $y \ni \varphi, \neg\psi$ to eliminate a problem $\neg(\varphi \rhd \psi) \in x \cap \mathcal{D}$, we did indeed eliminate it, as for no z with yS_xz we have $\psi \in z$. However, this should hold for any future expansion of the model too. Thus, any time we eliminate a problem $\neg(\varphi \rhd \psi) \in x \cap \mathcal{D}$, we introduce a world y with a promise that in no future time we will be able to go to a world z containing ψ via an S_x-transition. Somehow we should keep track of all these promises throughout the construction and make sure that all the promises are indeed kept. This is taken care of by our so called ψ-critical cones (see for example also [dJJ98]). As ψ is certainly not allowed to hold in R-successors of y, it is reasonable to demand that $\Box\neg\psi \in y$. (Where y was introduced to eliminate the problem $\neg(\varphi \rhd \psi) \in x \cap \mathcal{D}$.)

Note that problems have quantifier complexity $\exists\forall$. We have chosen to call them problems due to their prominent existential nature.

3.2 Some methods to obtain completeness

For modal logics in general, quite an arsenal of methods to obtain completeness is available. For instance the standard operations on canonical models like path–coding (unraveling), filtrations and bulldozing (see [BV01]). Or one can mention uniform methods like the use of Shalqvist formulas or the David Lewis theorem [Boo93]. A very secure method is to construct counter models piece by piece. A nice example can be found in [Boo93], Chapter 10. In [HV01] and in [HH02] a step-by-step method is exposed in the setting of universal algebras. New approximations of the model are given by moves in an (infinite) game.

For interpretability logics the available methods are rather limited in number. In the case of the basic logic **IL** a relatively simple unraveling works. Although **IL**M does allow a same treatment, the proof is already much less clear. (For both proofs, see [dJJ98]). However, for logics that contain **IL**M_0 but not $\mathbf{IL}\mathsf{M}$ it is completely unclear how to obtain completeness via an unraveling and we are forced into more secure methods like the above mentioned building of models piece by piece. And this is precisely what we do in this paper.

Decidability and the finite model property are two related issues that more or less seem to divide the landscape of interpretability logics into the same classes. That is, the proof that **IL** has the finite model property is relatively easy. The same can be said about $\mathbf{IL}\mathsf{M}$. For logics like $\mathbf{IL}\mathsf{M}_0$ the issue seems much more involved and a proper proof of the finite model property, if one exists at all, has not been given yet. Alternatively, one could resort to other methods for showing decidability like the Mosaic method [BV01].

4 The construction method

4.1 Preparing the construction

An $\mathbf{IL}\mathsf{X}$-labeled frame is just a Veltman frame in which every node is labeled by a maximal $\mathbf{IL}\mathsf{X}$-consistent set and some R-transitions are labeled by a formula. R-transitions labeled by a formula C indicate that some C-criticallity is essentially present at this place.

Definition 4.1. An $\mathbf{IL}\mathsf{X}$-*labeled frame* is a quadruple $\langle W, R, S, \nu\rangle$. Here $\langle W, R, S\rangle$ is an **IL**-frame and ν is a labeling function. The function ν assigns to each $x \in W$ a maximal $\mathbf{IL}\mathsf{X}$-consistent set of sentences $\nu(x)$. To some pairs $\langle x, y\rangle$ with xRy, ν assigns a formula $\nu(\langle x, y\rangle)$.

If there is no chance of confusion we will just speak of labeled frames or even just of frames rather than $\mathbf{IL}\mathsf{X}$-labeled frames. Labeled frames inherit all the terminology and notation from normal frames. Note that an $\mathbf{IL}\mathsf{X}$-labeled frame need not be, and shall in general not be, an $\mathbf{IL}\mathsf{X}$-frame. If we speak about a labeled $\mathbf{IL}\mathsf{X}$-frame we always mean an $\mathbf{IL}\mathsf{X}$-labeled $\mathbf{IL}\mathsf{X}$-frame. To indicate that $\nu(\langle x, y\rangle) = A$ we will sometimes write xR^Ay or $\nu(x, y) = A$.

Formally, given $F = \langle W, R, S, \nu\rangle$, one can see ν as a subset of $(W \cup (W \times W)) \times (\mathsf{Form}_{\mathbf{IL}} \cup \{\Gamma \mid \Gamma \text{ is a maximal } \mathbf{IL}\mathsf{X} \text{ consistent set}\})$ such that the following properties hold.

- $\forall\, x {\in} W\ (\langle x, y\rangle \in \nu \Rightarrow y \text{ is a MCS})$
- $\forall\, \langle x, y\rangle {\in} W \times W\ (\langle\langle x, y\rangle, z\rangle \in \nu \Rightarrow z \text{ is a formula})$
- $\forall\, x {\in} W \exists y\ \langle x, y\rangle \in \nu$
- $\forall x, y, y' (\langle x, y\rangle \in \nu \wedge \langle x, y'\rangle \in \nu \to y = y')$

We will often regard ν as a partial function on $W \cup (W \times W)$ which is total on W and which has its values in $\mathsf{Form}_{\mathbf{IL}} \cup \{\Gamma \mid \Gamma \text{ is a maximal } \mathbf{IL}\mathsf{X} \text{ consistent set}\}$

Remark 4.2. Every $\mathbf{IL}\mathsf{X}$-labeled frame $F = \langle W, R, S, \nu\rangle$ can be transformed to an **IL**-model $\overline{F}$ in a uniform way by defining for propositional variables p the valuation as $\overline{F}, x \Vdash p$ iff. $p \in \nu(x)$. By Corollary 2.17 we can also regard any model M satisfying the **IL**X frame condition[7] as an **IL**X-labeled frame $\overline{M}$ by defining $\nu(m) := \{\varphi \mid M, m \Vdash \varphi\}$.

We sometimes refer to $\overline{F}$ as the model induced by the frame F. Alternatively we will speak about the model corresponding to F. Note that for **IL**X-models M, we have $\overline{\overline{M}} = M$, but in general $\overline{\overline{F}} \neq F$ for **IL**X-labeled frames F.

Definition 4.3. Let x be a world in some **IL**X-labeled frame $\langle W, R, S, \nu\rangle$. The *$C$-critical cone above x*, we write $\mathcal{C}_x^C$, is defined inductively as

- $\nu(\langle x, y\rangle) = C \Rightarrow y \in \mathcal{C}_x^C$
- $x' \in \mathcal{C}_x^C$ & $x' S_x y \Rightarrow y \in \mathcal{C}_x^C$
- $x' \in \mathcal{C}_x^C$ & $x' R y \Rightarrow y \in \mathcal{C}_x^C$

Definition 4.4. Let x be a world in some **IL**X-labeled frame $\langle W, R, S, \nu\rangle$. The *generalized C-cone above x*, we write $\mathcal{G}_x^C$, is defined inductively as

- $y \in \mathcal{C}_x^C \Rightarrow y \in \mathcal{G}_x^C$
- $x' \in \mathcal{G}_x^C$ & $x' S_w z \Rightarrow z \in \mathcal{G}_x^C$ for arbitrary w
- $x' \in \mathcal{G}_x^C$ & $x' R y \Rightarrow y \in \mathcal{G}_x^C$

It follows directly from the definition that the C-critical cone above x is part of the generalized C-cone above x. So, if $\mathcal{G}_x^B \cap \mathcal{G}_x^C = \varnothing$, then certainly $\mathcal{C}_x^B \cap \mathcal{C}_x^C = \varnothing$.

We also note that there is some redundancy in Definitions 4.3 and 4.4. The last clause in the inductive definitions demands closure of the cone under R-successors. But from Definition 2.9.5 closure of the cone under R follows from closure of the cone under S_x. We have chosen to explicitly adopt the closure under the R. In doing so, we obtain a notion that serves us also in the environment of so-called quasi frames (see Definition 5.1) in which not necessarily $(x{\upharpoonright})^2 \cap R \subseteq S_x$.

Definition 4.5. Let $F = \langle W, R, S, \nu\rangle$ be a labeled frame and let $\overline{F}$ be the induced **IL**-model. Furthermore, let $\mathcal{D}$ be some set of sentences. We say that *a truth lemma holds in F with respect to $\mathcal{D}$* if $\forall\, A{\in}\mathcal{D}\ \forall\, x{\in}\overline{F}$

$$\overline{F}, x \Vdash A \Leftrightarrow A \in \nu(x).$$

If there is no chance of confusion we will omit some parameters and just say "a truth lemma holds at F" or even "a truth lemma holds". The following definitions give us a means to measure how far we are away from a truth lemma.

[7] We could even say, any **IL**X-model.

Definition 4.6 (Temporary definition). [8] Let $\mathcal{D}$ be some set of sentences and let $F = \langle W, R, S, \nu\rangle$ be an **IL**X-labeled frame. A *$\mathcal{D}$-problem* is a pair $\langle x, \neg(A \rhd B)\rangle$ such that $\neg(A \rhd B) \in \nu(x) \cap \mathcal{D}$ and for every y with xRy we have $[A \in \nu(y) \Rightarrow \exists z\ (yS_x z \wedge B \in \nu(z))]$.

Definition 4.7 (Deficiencies). Let $\mathcal{D}$ be some set of sentences and let $F = \langle W, R, S, \nu\rangle$ be an **IL**X-labeled frame. A *$\mathcal{D}$-deficiency* is a triple $\langle x, y, C \rhd D\rangle$ with xRy, $C \rhd D \in \nu(x) \cap \mathcal{D}$, and $C \in \nu(y)$, but for no z with $yS_x z$ we have $D \in \nu(z)$.

If the set $\mathcal{D}$ is clear or fixed, we will just speak about problems and deficiencies.

Definition 4.8. Let A be a formula. We define the *single negation* of A, we write $\sim\!A$, as follows. If A is of the form $\neg B$ we define $\sim\!A$ to be B. If A is not a negated formula we set $\sim\!A := \neg A$.

The next lemma shows that a truth lemma w.r.t. $\mathcal{D}$ can be reformulated in the combinatoric terms of deficiencies and problems. (See also the equivalence of $(*)$ and $(**)$ in Section 3.)

Lemma 4.9. *Let $F = \langle W, R, S, \nu\rangle$ be a labeled frame, and let $\mathcal{D}$ be a set of sentences closed under single negation and subformulas. A truth lemma holds in F w.r.t. $\mathcal{D}$ iff. there are no $\mathcal{D}$-problems nor $\mathcal{D}$-deficiencies.*

Proof. The proof is really very simple and precisely shows they interplay between all the ingredients. $\dashv$

The labeled frames we will construct are always supposed to satisfy some minimal reasonable requirements. We summarize these in the notion of adequacy.

Definition 4.10 (Adequate frames). A frame is called *adequate* if the following conditions are satisfied.

1. $xRy \Rightarrow \nu(x) \prec \nu(y)$
2. $A \neq B \Rightarrow \mathcal{G}_x^A \cap \mathcal{G}_x^B = \varnothing$
3. $y \in \mathcal{C}_x^A \Rightarrow \nu(x) \prec_A \nu(y)$

If no confusion is possible we will just speak of frames instead of adequate labeled frames. As a matter of fact, all the labeled frames we will see from now on will be adequate. In the light of adequacy it seems reasonable to work with a slightly more elegant definition of a $\mathcal{D}$-problem.

Definition 4.11 (Problems). Let $\mathcal{D}$ be some set of sentences. A *$\mathcal{D}$-problem* is a pair $\langle x, \neg(A \rhd B)\rangle$ such that $\neg(A \rhd B) \in \nu(x) \cap \mathcal{D}$ and for no $y \in \mathcal{C}_x^B$ we have $A \in \nu(y)$.

[8] We will eventually work with Definition 4.11.

From now on, this will be our working definition. Clearly, on adequate labeled frames, if $\langle x, \neg(A \rhd B)\rangle$ is not a problem in the new sense, it is not a problem in the old sense.

Remark 4.12. It is also easy to see that the we still have the interesting half of Lemma 4.9. Thus, we still have, that a truth lemma holds if there are no deficiencies nor problems.

To get a truth lemma we have to somehow get rid of problems and deficiencies. This will be done by adding bits and pieces to the original labeled frame. Thus the notion of an extension comes into play.

Definition 4.13 (Extension). Let $F = \langle W, R, S, \nu\rangle$ be a labeled frame. We say that $F' = \langle W', R', S', \nu'\rangle$ is an *extension* of F, we write $F \subseteq F'$, if $W \subseteq W'$ and the relations in F' restricted to F yield the corresponding relations in F.

More formally, the requirements on the restrictions in the above definition amount to saying that for $x, y, z \in F$ we have the following.

- $xR'y$ iff. xRy
- $yS'_x z$ iff. $yS_x z$
- $\nu'(x) = \nu(x)$
- $\nu'(\langle x, y\rangle)$ is defined iff. $\nu(\langle x, y\rangle)$ is defined, and in this case $\nu'(\langle x, y\rangle) = \nu(\langle x, y\rangle)$.

A problem in F is said to be *eliminated* by the extension F' if it is no longer a problem in F'. Likewise we can speak about elimination of deficiencies.

Definition 4.14 (Depth). The *depth* of a finite frame F, we will write $\mathsf{depth}(F)$ is the maximal length of sequences of the form $x_0 R \ldots R x_n$. (For convenience we define $\max(\varnothing) = 0$.)

Definition 4.15 (Union of Bounded Chains). An indexed set $\{F_i\}_{i\in\omega}$ of labeled frames is called a *chain* if for all i, $F_i \subseteq F_{i+1}$. It is called a *bounded chain* if for some number n, $\mathsf{depth}(F_i) \leq n$ for all $i \in \omega$. The *union* of a bounded chain $\{F_i\}_{i\in\omega}$ of labeled frames F_i is defined as follows.

$$\cup_{i\in\omega} F_i := \langle \cup_{i\in\omega} W_i, \cup_{i\in\omega} R_i, \cup_{i\in\omega} S_i, \cup_{i\in\omega} \nu_i\rangle$$

It is clear why we really need the boundedness condition. We want the union to be an **IL**-frame. So, certainly R should be conversely well-founded. This can only be the case if our chain is bounded.

4.2 The main lemma

We now come to the main motor behind many results. It is formulated in rather general terms so that it has a wide range of applicability. As a draw-back, we get that any application still requires quite some work.

Lemma 4.16 (Main Lemma). *Let* $\mathbf{IL}\mathsf{X}$ *be an interpretability logic and let* $\mathcal{C}$ *be a (first or higher order) frame condition such that for any* $\mathbf{IL}$*-frame* F *we have*

$$F \models \mathcal{C} \Rightarrow F \models \mathsf{X}.$$

Let $\mathcal{D}$ *be a finite set of sentences. Let* $\mathcal{I}$ *be a set of so-called* invariants *of labeled frames so that we have the following properties.*

- $F \models \mathcal{I}^{\mathcal{U}} \Rightarrow F \models \mathcal{C}$*, where* $\mathcal{I}^{\mathcal{U}}$ *is that part of* $\mathcal{I}$ *that is closed under bounded unions of labeled frames.*
- $\mathcal{I}$ *contains the following invariant:* $xRy \rightarrow \exists\, A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box\neg D \mid D$ *a subformula of some* $B \in \mathcal{D}\}$.
- *For any adequate labeled frame* F*, satisfying all the invariants, we have the following.*
  - *Any* $\mathcal{D}$*-problem of* F *can be eliminated by extending* F *in a way that conserves all invariants.*
  - *Any* $\mathcal{D}$*-deficiency of* F *can be eliminated by extending* F *in a way that conserves all invariants.*

In case such a set of invariants $\mathcal{I}$ *exists, we have that any* $\mathbf{IL}\mathsf{X}$*-labeled adequate frame* F *satisfying all the invariants can be extended to some labeled adequate* $\mathbf{IL}\mathsf{X}$*-frame* $\hat{F}$ *on which a truth-lemma with respect to* $\mathcal{D}$ *holds.*

Moreover, if for any finite $\mathcal{D}$ *that is closed under subformulas and single negations, a corresponding set of invariants* $\mathcal{I}$ *can be found as above and such that moreover* $\mathcal{I}$ *holds on any one-point labeled frame, we have that* $\mathbf{IL}\mathsf{X}$ *is a complete logic.*

Proof. By subsequently eliminating problems and deficiencies by means of extensions. These elimination processes have to be robust in the sense that every problem or deficiency that has been dealt with, should not possibly re-emerge. But, the requirements of the lemma almost immediately imply this.

For the second part of the Main Lemma, we suppose that for any finite set $\mathcal{D}$ closed under subformulas and single negations, we can find a corresponding set of invariants $\mathcal{I}$. If now, for any such $\mathcal{D}$, all the corresponding invariants $\mathcal{I}$ hold on any one-point labeled frame, we are to see that $\mathbf{IL}\mathsf{X}$ is a complete logic, that is, $\mathbf{IL}\mathsf{X} \nvdash A \Rightarrow \exists M\ (M \models X \ \&\ M \models \neg A)$.

But this just follows from the above. If $\mathbf{IL}\mathsf{X} \nvdash A$, we can find a maximal $\mathbf{IL}\mathsf{X}$-consistent set Γ with $\neg A \in \Gamma$. Let $\mathcal{D}$ be the smallest set that contains $\neg A$ and is closed under subformulas and single negations and consider the invariants corresponding to $\mathcal{D}$. The labeled frame $F := \langle \{x\}, \varnothing, \varnothing, \langle x, \Gamma\rangle\rangle$ can thus be extended to a labeled adequate $\mathbf{IL}\mathsf{X}$-frame $\hat{F}$ on which a truth lemma with respect to $\mathcal{D}$ holds. Thus certainly $\overline{\hat{F}}, x \Vdash \neg A$, that is, A is not valid on the model induced by $\hat{F}$. $\dashv$

The construction method can also be used to obtain decidability via the finite model property. In such a case, one should re-use worlds that were introduced earlier in the construction.

The following two lemmata indicate how good labels can be found for the elimination of problems and deficiencies.

Lemma 4.17. *Let Γ be a maximal* **IL**X*-consistent set such that* $\neg(A \rhd B) \in \Gamma$. *Then there exists a maximal* **IL**X*-consistent set* Δ *such that* $\Gamma \prec_B \Delta \ni A, \Box\neg A$.

Proof. So, consider $\neg(A \rhd B) \in \Gamma$, and suppose that no required Δ exists. We can then find a[9] formula C for which $C \rhd B \in \Gamma$ such that

$$\neg C, \Box\neg C, A, \Box\neg A \vdash_{\mathbf{IL}\mathsf{X}} \bot.$$

Consequently

$$\vdash_{\mathbf{IL}\mathsf{X}} A \wedge \Box\neg A \to C \vee \Diamond C$$

and thus, by Lemma 2.2, also $\vdash_{\mathbf{IL}\mathsf{X}} A \rhd C$. But as $C \rhd B \in \Gamma$, also $A \rhd B \in \Gamma$. This clearly contradicts the consistency of Γ. ⊣

For deficiencies there is a similar lemma.

Lemma 4.18. *Consider* $C \rhd D \in \Gamma \prec_B \Delta \ni C$. *There exists* Δ' *with* $\Gamma \prec_B \Delta' \ni D, \Box\neg D$.

Proof. Suppose for a contradiction that $C \rhd D \in \Gamma \prec_B \Delta \ni C$ and there does not exist a Δ' with $\Gamma \prec_B \Delta' \ni D, \Box\neg D$. Taking the contraposition of Lemma 4.17 we get that $\neg(D \rhd B) \notin \Gamma$, whence $D \rhd B \in \Gamma$ and also $C \rhd B \in \Gamma$. This clearly contradicts the consistency of Δ as $\Gamma \prec_B \Delta \ni C$. ⊣

4.3 Completeness and the main lemma

The main lemma provides a powerful method for proving modal completeness. In several cases it is actually the only known method available.

Remark 4.19. A modal completeness proof for an interpretability logic **IL**X is by the main lemma reduced to the following four ingredients.

- **Frame Condition** Providing a frame condition $\mathcal{C}$ and a proof that

$$F \models \mathcal{C} \Rightarrow F \models \mathbf{IL}\mathsf{X}.$$

- **Invariants** Given a finite set of sentences $\mathcal{D}$ (closed under subformulas and single negations), providing invariants $\mathcal{I}$ that hold for any one-point labeled frame. Certainly $\mathcal{I}$ should contain $xRy \to \exists\, A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}$.

[9] Writing out the definition and by compactness, we get a finite number of formulas $C_1, \ldots, C_n$ with $C_i \rhd B \in \Gamma$, such that $\neg C_1, \ldots, \neg C_n, \Box\neg C_1, \ldots, \Box\neg C_n, A, \Box\neg A \vdash_{\mathbf{IL}\mathsf{X}} \bot$. We can now take $C := C_1 \vee \ldots \vee C_n$. Clearly, as all the $C_i \rhd B \in \Gamma$, also $C \rhd B \in \Gamma$.

- **elimination**
  - **Problems** Providing a procedure of elimination by extension for problems in labeled frames that satisfy all the invariants. This procedure should come with a proof that it preserves all the invariants.
  - **Deficiencies** Providing a procedure of elimination by extension for deficiencies in labeled frames that satisfy all the invariants. Also this procedure should come with a proof that it preserves all the invariants.
- **Rounding up** A proof that for any bounded chain of labeled frames that satisfy the invariants, automatically, the union satisfies the frame condition $\mathcal{C}$ of the logic.

The completeness proofs that we will present will all have the same structure, also in their preparations. As we will see, eliminating problems is more elementary than eliminating deficiencies.

As we already pointed out, we eliminate a problem by adding some new world plus an adequate label to the model we had. Like this, we get a structure that need not even be an **IL**-model. For example, in general, the R relation is not transitive. To come back to at least an **IL**-model, we should close off the new structure under transitivity of R and S et cetera. This closing off is in its self an easy and elementary process but we do want that the invariants are preserved under this process. Therefore we should have started already with a structure that admitted a closure. Actually in this paper we will always want to obtain a model that satisfies the frame condition of the logic.

The preparations to a completeness proof in this paper thus have the following structure.

- Determining a frame condition for **IL**X and a corresponding notion of an **IL**X-frame.
- Defining a notion of a quasi **IL**X-frame.
- Defining some notions that remain constant throughout the closing of quasi **IL**X-frames, but somehow capture the dynamic features of this process.
- Proving that a quasi **IL**X-frame can be closed off to an adequate labeled **IL**X-frame.
- Preparing the elimination of deficiencies.

The most difficult job in a the completeness proofs we present in this paper, was in finding correct invariants and in preparing the elimination of deficiencies. Once this is fixed, the rest follows in a rather mechanical way. Especially the closure of quasi **IL**X-frames to adequate **IL**X-frames is a very laborious enterprise.

5 The logic IL

The modal logic **IL** has been proved to be modally complete in [dJV90]. We shall reprove the completeness here using the main lemma.

The completeness proof of **IL** can be seen as the mother of all our completeness proofs in interpretability logics. Not only does it reflect the general structure of applications of the Main Lemma clearly, also it so that we can use large parts of the preparations to the completeness proof of **IL** in other proofs too. Especially closability proofs are cumulative. Thus, we can use the lemma that any quasi-frame is closable to an adequate frame, in any other completeness proof.

5.1 Preparations

Definition 5.1. A *quasi-frame* G is a quadruple $\langle W, R, S, \nu\rangle$. Here W is a non-empty set of worlds, and R a binary relation on W. S is a set of binary relations on W indexed by elements of W. The ν is a labeling as defined on labeled frames. Critical cones and generalized cones are defined just in the same way as in the case of labeled frames. G should posess the following properties.

1. R is conversely well-founded
2. $yS_xz \rightarrow xRy$ & xRz
3. $xRy \rightarrow \nu(x) \prec \nu(y)$
4. $A \neq B \rightarrow \mathcal{G}_x^A \cap \mathcal{G}_x^B = \varnothing$
5. $y \in \mathcal{C}_x^A \rightarrow \nu(x) \prec_A \nu(y)$

Clearly, adequate labeled frames are special cases of quasi frames. Quasi-frames inherit all the notations from labeled frames. In particular we can thus speak of chains and the like.

Lemma 5.2 (IL-closure). *Let $G = \langle W, R, S, \nu\rangle$ be a quasi-frame. There is an adequate* **IL***-frame F extending G. That is, $F = \langle W, R', S', \nu\rangle$ with $R \subseteq R'$ and $S \subseteq S'$.*

Proof. We define an *imperfection* on a quasi-frame F_n to be a tuple γ having one of the following forms.

(*i*) $\gamma = \langle 0, a, b, c\rangle$ with $F_n \models aRbRc$ but $F_n \not\models aRc$

(*ii*) $\gamma = \langle 1, a, b\rangle$ with $F_n \models aRb$ but $F_n \not\models bS_ab$

(*iii*) $\gamma = \langle 2, a, b, c, d\rangle$ with $F_n \models bS_acS_ad$ but not $F_n \models bS_ad$

(*iv*) $\gamma = \langle 3, a, b, c\rangle$ with $F_n \models aRbRc$ but $F_n \not\models bS_ac$

Now let us start with a quasi-frame $G = \langle W, R, S, \nu \rangle$. We will define a chain of quasi-frames. Every new element in the chain will have at least one imperfection less than its predecessor. The union will have no imperfections at all. It will be our required adequate **IL**-frame.

Let $<_0$ be the well-ordering on

$$C := (\{0\} \times W^3) \cup (\{1\} \times W^2) \cup (\{2\} \times W^4) \cup (\{3\} \times W^3)$$

induced by the occurrence order in some fixed enumeration of C. We define our chain to start with

$F_0 := G$. To go from F_n to F_{n+1} we proceed as follows. Let γ be the $<_0$-minimal imperfection on F_n. In case no such γ exists we set $F_{n+1} := F_n$. If such a γ does exist, F_{n+1} is as dicted by the case distinctions.

(i) $F_{n+1} := \langle W_n, R_n \cup \{\langle a, c\rangle\}, S_n, \nu_n \rangle$

(ii) $F_{n+1} := \langle W_n, R_n, S_n \cup \{\langle a, b, b\rangle\}, \nu_n \rangle$

(iii) $F_{n+1} := \langle W_n, R_n, S_n \cup \{\langle a, b, d\rangle\}, \nu_n \rangle$

(iv) $F_{n+1} := \langle W_n, R_n \cup \{\langle a, c\rangle\}, S_n \cup \{\langle a, b, c\rangle\}, \nu_n \rangle$

By an easy but elaborate induction, we can see that each F_n is a quasi-frame. The induction boils down to checking for each case (i)-(iv) that all the properties (1)-(5) from Definition 5.1 remain valid.

Instead of proving (4) and (5), it is better to prove something stronger, that is, that the critical and generalized cones remain unchanged.

4'. $\forall n\ [F_{n+1} \models y \in \mathcal{G}_x^A \Leftrightarrow F_n \models y \in \mathcal{G}_x^A]$

5'. $\forall n\ [F_{n+1} \models y \in \mathcal{C}_x^A \Leftrightarrow F_n \models y \in \mathcal{C}_x^A]$

Next, it is not hard to prove that $F := \cup_{i\in\omega} F_i$ is the required adequate **IL**-frame. To this extent, the following properties have to be checked. All properties have easy proofs.

$(a.)$ W is the domain of F	$(g.)$ $F \models xRy \rightarrow yS_xy$
$(b.)$ $R_0 \subseteq \cup_{i\in\omega} R_i$	$(h.)$ $F \models xRyRz \rightarrow yS_xz$
$(c.)$ $S_0 \subseteq \cup_{i\in\omega} S_i$	$(i.)$ $F \models uS_xvS_xw \rightarrow uS_xw$
$(d.)$ R is conv. wellfounded on F	$(j.)$ $F \models xRy \Rightarrow \nu(x) \prec \nu(y)$
$(e.)$ $F \models xRyRz \rightarrow xRz$	$(k.)$ $A \neq B \Rightarrow F \models \mathcal{G}_x^A \cap \mathcal{G}_x^B = \varnothing$
$(f.)$ $F \models yS_xz \rightarrow xRy\ \&\ xRz$	$(l.)$ $F \models y \in \mathcal{C}_x^A \Rightarrow \nu(x) \prec_A \nu(y)$

$\dashv$

We note that the **IL**-frame $F \supseteq G$ from above is actually the minimal one extending G. If in the sequel, if we refer to the closure given by the lemma, we shall mean this minimal one. Also do we note that the proof is independent on the enumeration of C and hence the order $<_0$ on C. The lemma can also be applied to non-labeled structures. If we drop all the requirements on the labels in Definition 5.1 and in Lemma 5.2 we end up with a true statement about just the old **IL**-frames.

Lemma 5.2 also allows a very short proof running as follows. Any intersection of adequate **IL**-frames with the same domain is again an adequate **IL**-frame. There is an adequate **IL**-frame extending G. Thus by taking intersections we find a minimal one. We have chosen to present our explicit proof as they allow us, now and in the sequel, to see which properties remain invariant.

Corollary 5.3. *Let $\mathcal{D}$ be a finite set of sentences, closed under subformulas and single negations. Let $G = \langle W, R, S, \nu\rangle$ be a quasi-frame on which*

$$xRy \to \exists\, A{\in}((\nu(y) \setminus \nu x) \cap \{\Box D \mid D \in \mathcal{D}\}) \quad (*)$$

holds. Property $(*)$ does also hold on the **IL***-closure F of G.*

Proof. We can just take the property along in the proof of Lemma 5.2. In Case (i) and (iv) we note that $aRbRc \to \nu(a) \subseteq_{\Box} \nu(c)$. Thus, if $A{\in}((\nu(c) \setminus \nu(b)) \cap \{\Box D \mid D \in \mathcal{D}\})$, then certainly $A \notin \nu(a)$. $\dashv$

We have now done all the preparations for the completeness proof. Normally, also a lemma is needed to deal with deficiencies. But in the case of **IL**, Lemma 4.18 suffices.

5.2 Modal completeness

Theorem 5.4. **IL** *is a complete logic*

Proof. We specify the four ingredients mentioned in Remark 4.19.

Frame Condition For **IL**, the frame condition is empty, that is, every frame is an **IL** frame.

Invariants Given a finite set of sentences $\mathcal{D}$ closed under subformulas and single negation, the only invariant is $xRy \to \exists\, A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}$. Clearly this invariant holds on any one-point labeled frame.

Elimination So, let $F := \langle W, R, S, \nu\rangle$ be a labeled frame satisfying the invariant. We will see how to eliminate both problems and deficiencies while conserving the invariant.

Problems Any problem $\langle a, \neg(A \rhd B)\rangle$ of F will be eliminated in two steps.

1. With Lemma 4.17 we find Δ with $\nu(a) \prec_B \Delta \ni A, \Box\neg A$. We fix some $b \notin W$. We now define

$$G' := \langle W \cup \{b\}, R \cup \{\langle a, b\rangle\}, S, \nu \cup \{\langle b, \Delta\rangle, \langle\langle a, b\rangle, B\rangle\}\rangle.$$

   It is easy to see that G' is actually a quasi-frame. Note that if $G' \models xRb$, then x must be a and whence $\nu(x) \prec \nu(b)$ by definition of $\nu(b)$. Also it is not hard to see that if $b \in \mathcal{C}_x^C$ for $x{\neq}a$, that then $\nu(x) \prec_C \nu(b)$. For, $b \in \mathcal{C}_x^C$ implies $a \in \mathcal{C}_x^C$ whence $\nu(x) \prec_C \nu(a)$. By $\nu(a) \prec \nu(b)$ we get that $\nu(x) \prec_C \nu(b)$. In case $x{=}a$ we see that by definition $b \in \mathcal{C}_a^B$. But, we have chosen Δ so that $\nu(a) \prec_B \nu(b)$. We also see that G' satisfies the invariant as $\Box\neg A \in \nu(b) \setminus \nu(a)$ and $\sim A \in \mathcal{D}$.

2. With Lemma 5.2 we extend G' to an adequate labeled **IL**-frame G. Corollary 5.3 tells us that the invariant indeed holds at G. Clearly $\langle a, \neg(A \rhd B)\rangle$ is no longer a problem in G.

Deficiencies. Again, any deficiency $\langle a, b, C \rhd D\rangle$ in F will be eliminated in two steps.

1. We first define B to be the formula such that $b \in \mathcal{C}_a^B$. If such a B does not exist, we take B to be $\bot$. Note that if such a B does exist, it must be unique by Property 4 of Definition 5.1. By Lemma 4.18 we can now find a Δ' such that $\nu(a) \prec_B \Delta' \ni D, \Box\neg D$. We fix some $c \notin W$ and define

$$G' := \langle W, R \cup \{a, c\}, S \cup \{a, b, c\}, \nu \cup \{c, \Delta'\}\rangle.$$

Again it is not hard to see that G' is a quasi-frame that satisfies the invariant. Clearly R is conversely well-founded. The only new S in G' is bS_ac, but we also defined aRc. We have chosen Δ' such that $\nu(a) \prec_B \nu(c)$. Clearly $\Box\neg D \notin \nu(a)$.

2. Again, G' is closed off under the frame conditions with Lemma 5.2. Again we note that the invariant is preserved in this process. Clearly $\langle a, b, C \rhd D\rangle$ is not a deficiency in G.

Rounding up Clearly the union of a bounded chain of **IL**-frames is again an **IL**-frame.

$\dashv$

It is well known that **IL** has the finite model property and whence is decidable. With some more effort however we could have obtained the finite model property using the Main Lemma. We have chosen not to do so, as for our purposes the completeness via the construction method is sufficient.

Also, to obtain the finite model property, one has to re-use worlds during the construction method. The constraints on which worlds can be re-used is per logic differently. One aim of this section was to prove some results on a construction that is present in all other completeness proofs too. Therefore we needed some uniformity and did not want to consider re-using of worlds.

6 The Logic ILM

The modal completeness of **IL**M was proved by de Jongh and Veltman in [dJV90]. In this section we will reprove the modal completeness of the logic

ILM via the Main Lemma. The general approach is not much different from the completeness proof for **IL**.

The novelty consists of incorporating the **IL**M frame condition, that is, whenever yS_xzRu holds, we should also have yRu. In this case, adequacy imposes $\nu(y) \prec \nu(u)$.

Thus, whenever we introduce an S_x relation, when eliminating a deficiency, we should keep in mind that in a later stage, this S_x can activate the **IL**M frame condition. It turns out to be sufficient to demand $\nu(y) \subseteq_\Box \nu(z)$ whenever ySz. Also, we should do some additional book keeping as to keep our critical cones fit to our purposes.

6.1 Preparations

Let us first recall the principle M, also called Montagna's principle.

$$\mathsf{M}: \quad A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C$$

Definition 6.1. An **IL**M-frame is a frame such that $yS_xzRu \to yRu$ holds on it. A(n adequate) labeled **IL**M-frame is an adequate labeled **IL**M-frame on which $yS_xz \to \nu(y) \subseteq_\Box \nu(z)$ holds. We call $yS_xzRu \to yRu$ the frame condition of **IL**M.

The next lemma tells us that the frame condition of **IL**M, indeed characterizes the frames of **IL**M.

Lemma 6.2. $F \models \forall x, y, u, v\ (yS_xuRv \to yRv) \Leftrightarrow F \models \mathbf{IL}\mathsf{M}$

We will now introduce a notion of a quasi-**IL**M-frame and a corresponding closure lemma. In order to get an **IL**M-closure lemma in analogy with Lemma 5.2 we need to introduce a technicality.

Definition 6.3. The A-critical $\mathcal{M}$-cone of x, we write $\mathcal{M}^A_x$, is defined inductively as follows.

- $xR^Ay \to y \in \mathcal{M}^A_x$
- $y \in \mathcal{M}^A_x \ \&\ yRz \to z \in \mathcal{M}^A_x$
- $y \in \mathcal{M}^A_x \ \&\ yS_xz \to z \in \mathcal{M}^A_x$
- $y \in \mathcal{M}^A_x \ \&\ yS^{\mathsf{tr}}uRv \to v \in \mathcal{M}^A_x$

Definition 6.4. A quasi-frame is a quasi-**IL**M-frame if[10] the following properties hold.

[10]By R^{tr} we denote the transitive closure of R, inductively defined as the smallest set such that $xRy \to xR^{\mathsf{tr}}y$ and $\exists z\ (xR^{\mathsf{tr}}z \wedge zR^{\mathsf{tr}}y) \to xR^{\mathsf{tr}}y)$. Similarly we define S^{tr}. The ; is the composition operator on relations. Thus, for example, $y(R^{\mathsf{tr}};S)z$ iff. there is a u such that $yR^{\mathsf{tr}}u$ and uSz. Recall that uSv iff. uS_xv for some x. In the literature one often also uses the $\circ$ notation, where $xR \circ Sy$ iff. $\exists z\ xSzRy$. Note that $R^{\mathsf{tr}}; S^{\mathsf{tr}}$ is conversely well-founded iff. $R^{\mathsf{tr}} \circ S^{\mathsf{tr}}$ is conversely well-founded.

- $R^{\mathsf{tr}}; S^{\mathsf{tr}}$ is conversely well-founded[11]
- $yS_x z \to \nu(y) \subseteq_\Box \nu(z)$
- $y \in \mathcal{M}_x^A \Rightarrow \nu(x) \prec_A \nu(y)$

It is easy to see that $\mathcal{C}_x^A \subseteq \mathcal{M}_x^A \subseteq \mathcal{G}_x^A$. Thus we have that $A \neq B \to \mathcal{M}_x^A \cap \mathcal{M}_x^B = \varnothing$. Also, it is clear that if F is an **IL**M-frame, then $F \models \mathcal{M}_x^A = \mathcal{C}_x^A$. Actually we have that a quasi-**IL**M-frame F is an **IL**M-frame iff. $F \models \mathcal{M}_x^A = \mathcal{C}_x^A$.

Lemma 6.5 (ILM-closure). *Let $G = \langle W, R, S, \nu \rangle$ be a quasi-**IL**M-frame. There is an adequate **IL**M-frame F extending G. That is, $F = \langle W, R', S', \nu \rangle$ with $R \subseteq R'$ and $S \subseteq S'$.*

Proof. The proof is very similar to that of Lemma 5.2. As a matter of fact, we will use large parts of the latter proof in here. For quasi-**IL**M-frames we also define the notion of an imperfection. An *imperfection* on a quasi-**IL**M-frame F_n is a tuple γ that is either an imperfection on the quasi-frame F_n, or it is a tuple of the form

$$\gamma = \langle 4, a, b, c, d \rangle \text{ with } F_n \models bS_a cRd \text{ but } F_n \not\models bRd.$$

As in the closure proof for quasi-frames, we define a chain of quasi-**IL**M-frames. Each new frame in the chain will have at least one imperfection less than its predecessor. We only have to consider the new imperfections, in which case we define

$$F_{n+1} := \langle W_n, R_n \cup \{\langle b, d \rangle\}, S_n, \nu_n \rangle.$$

We now see by an easy but elaborate induction that every F_n is a quasi-**IL**M-frame. Again, this boils down to checking that at each of (i)-(v), all the eight properties from Definition 6.4 are preserved.

During the closure process, the critical cones do change. However, the critical $\mathcal{M}$-cones are invariant. Thus, it is useful to prove

$$8'. \ \ F_{n+1} \models y \in \mathcal{M}_x^A \text{ iff. } F_n \models y \in \mathcal{M}_x^A.$$

Our induction is completely straightforward. As an example we shall see that $8'$ holds in Case (i): We have eliminated an imperfection concerning the transitivity of the R relation and $F_{n+1} := \langle W_n, R_n \cup \{\langle a, c \rangle\}, S_n, \nu_n \rangle$.

To see that $8'$ holds, we reason as follows. Suppose $F_{n+1} \models y \in \mathcal{M}_x^A$. Thus $\exists z_1, \ldots, z_l$ $(0 \leq l)$ with[12] $F_{n+1} \models xR^A z_1(S_x \cup R \cup (S^{\mathsf{tr}}; R))z_2, \ldots, z_l(S_x \cup R \cup (S^{\mathsf{tr}}; R))y$. We transform the sequence $z_1, \ldots, z_l$ into a sequence $u_1, \ldots, u_m$

[11] In the case of quasi-frames we did not need a second order frame condition. We could use the second order frame condition of **IL** via $yS_xz \to xRy$ & xRz. Such a trick seems not to be available here.

[12] The union operator on relations can just be seen as the set-theoretical union. Thus, for example, $y(S_x \cup R)z$ iff. yS_xz or yRz.

$(0 \leq m)$ in the following way. Every occurrence of aRc in $z_1, \dots, z_l$ is replaced by $aRbRc$. In case that for some $n < l$ we have $z_n S^{\mathsf{tr}} aRc = z_{n+1}$, we replace z_n, z_{n+1} by z_n, b, c and thus $z_n(S^{\mathsf{tr}}; R)bRc$. We leave the rest of the sequence $z_1, \dots, z_l$ unchanged. Clearly $F_n \models xR^A u_1(S_x \cup R \cup (S^{\mathsf{tr}}; R))u_2, \dots, u_m(S_x \cup R \cup (S^{\mathsf{tr}}; R))y$, whence $F_n \models y \in \mathcal{M}_x^A$.

We shall include one more example for Case (v): We have eliminated an imperfection concerning the $\mathbf{IL}\mathsf{M}$ frame-condition and $F_{n+1} := \langle W_n, R_n \cup \{\langle b, d\rangle\}, S_n, \nu_n\rangle$. To see the conversely well-foundedness of R, we reason as follows. Suppose for a contradiction that there is an infinite sequence such that $F_{n+1} \models x_1 R x_2 R \dots$. We now get an infinite sequence $y_1, y_2, \dots$ by replacing every occurrence of bRd in $x_1, x_2, \dots$ by $bS_a cRd$ and leaving the rest unchanged. If there are infinitely many S_a-transitions in the sequence $y_1, y_2, \dots$ (note that there are certainly infinitely many R-transitions in $y_1, y_2, \dots$), we get a contradiction with our assumption that $R^{\mathsf{tr}}; S^{\mathsf{tr}}$ is conversely well-founded on F_n. In the other case we get a contradiction with the conversely well-foundedness of R on F_n.

Once we have seen that indeed, every F_n is a quasi-$\mathbf{IL}\mathsf{M}$-frame, it is not hard to see that $F := \cup_{i \in \omega} F_i$ is the required adequate $\mathbf{IL}\mathsf{M}$-frame. To this extend we have to check a list of properties $(a.)$-$(n.)$. The properties $(a.)$-$(l.)$ are as in the proof of Lemma 5.2.

The one exception is Property $(d.)$. To see $(d.)$, the conversely well-foundedness of R, we prove by induction on n that $F_n \models xRy$ iff. $F_0 \models x(S^{\mathsf{tr,refl}}; R^{\mathsf{tr}})y$. Thus, a hypothetical infinite sequence $F \models x_0 R x_1 R x_2 R \dots$ defines an infinite sequence $F_0 \models x_0(S^{\mathsf{tr,refl}}; R^{\mathsf{tr}})x_1(S^{\mathsf{tr,refl}}; R^{\mathsf{tr}})x_2 \dots$, which contradicts either the conversely well-foundedness of R or of $S^{\mathsf{tr}}; R^{\mathsf{tr}}$ on F_0.

The only new properties in this list are $(m.)$: $uS_x vRw \rightarrow uRw$ and $(n.)$: $yS_x z \rightarrow \nu(y) \subseteq_\Box \nu(z)$, but they are easily seen to hold on F. ⊣

Again do we note that the closure obtained in Lemma 6.5 is unique. Thus we can refer to the $\mathbf{IL}\mathsf{M}$-closure of a quasi-$\mathbf{IL}\mathsf{M}$-frame. All the information about the labels can be dropped in Definition 6.4 and Lemma 6.5 to obtain a lemma about regular $\mathbf{IL}\mathsf{M}$-frames.

Corollary 6.6. *Let $\mathcal{D}$ be a finite set of sentences, closed under subformulas and single negations. Let $G = \langle W, R, S, \nu\rangle$ be a quasi-$\mathbf{IL}\mathsf{M}$-frame on which*

$$xRy \rightarrow \exists\, A \in ((\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}) \quad (*)$$

holds. Property $(*)$ does also hold on the $\mathbf{IL}$-closure F of G.

Proof. The proof is as the proof of Corollary 5.3. We only need to remark on Case (v): If $bS_a cRd$, we have $\nu(b) \subseteq_\Box \nu(c)$. Thus, $A \in ((\nu(d) \setminus \nu(c)) \cap \{\Box D \mid D \in \mathcal{D}\})$ implies $A \notin \nu(b)$. ⊣

The final lemma in our preparations is a lemma that is needed to eliminate deficiencies properly.

Lemma 6.7. *Let Γ and Δ be maximal* **IL**M*-consistent sets. Consider $C \rhd D \in \Gamma \prec_B \Delta \ni C$. There exists a maximal* **IL**M*-consistent set Δ' with $\Gamma \prec_B \Delta' \ni D, \Box\neg D$ and $\Delta \subseteq_\Box \Delta'$.*

Proof. By compactness and by commutation of boxes and conjunctions, it is sufficient to show that for any formula $\Box E \in \Delta$ there is a Δ'' with $\Gamma \prec_B \Delta'' \ni D \wedge \Box E \wedge \Box\neg D$. As $C \rhd D$ is in the maximal **IL**M-consistent set Γ, also $C \wedge \Box E \rhd D \wedge \Box E \in \Gamma$. Clearly $C \wedge \Box E \in \Delta$, whence, by Lemma 4.18 we find a Δ'' with $\Gamma \prec_B \Delta'' \ni D \wedge \Box E \wedge \Box(\neg D \vee \neg \Box E)$. As **IL**M $\vdash \Box E \wedge \Box(\neg D \vee \neg \Box E) \to \Box\neg D$, we see that also $D \wedge \Box E \wedge \Box\neg D \in \Delta''$. ⊣

6.2 Completeness

Theorem 6.8. **IL**M *is a complete logic.*

Proof. **Frame Condition** In the case of **IL**M the frame condition is easy and well known, as expressed in Lemma 6.2.

Invariants Let $\mathcal{D}$ be a finite set of sentences closed under subformulas and single negations. We define a corresponding set of invariants.

$$\mathcal{I} := \begin{cases} xRy \to \exists\, A{\in}((\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}) \\ uS_x vRw \to uRw \end{cases}$$

Elimination Thus, we consider an **IL**M-labeled frame $F := \langle W, R, S, \nu\rangle$ that satisfies the invariants.

Problems Any problem $\langle a, \neg(A \rhd B)\rangle$ of F will be eliminated in two steps.

1. Using Lemma 4.17 we can find a MCS Δ with $\nu(a) \prec_B \Delta \ni A, \Box\neg A$. We fix some $b \notin W$ and define

   $$G' := \langle W \cup \{b\}, R \cup \{\langle a, b\rangle\}, S, \nu \cup \{\langle b, \Delta\rangle, \langle\langle a, b\rangle, B\rangle\}\rangle.$$

   We now see that G' is a quasi-**IL**M-frame. Thus, we need to check the eight points from Definitions 6.4 and 5.1. We will comment on some of these points.

   To see, for example, Point 4, $C \neq D \to \mathcal{G}^C_x \cap \mathcal{G}^D_x = \varnothing$, we reason as follows. First, we notice that $\forall\, x, y{\in}W\ [G' \models y \in \mathcal{G}^C_x$ iff. $F \models y \in \mathcal{G}^C_x]$ holds for any C. Suppose $G' \models \mathcal{G}^C_x \cap \mathcal{G}^D_x \neq \varnothing$. If $G' \models b \notin \mathcal{G}^C_x \cap \mathcal{G}^D_x$, then also $F \models \mathcal{G}^C_x \cap \mathcal{G}^D_x \neq \varnothing$. As F is an **IL**M-frame, it is certainly a quasi-**IL**M-frame, whence $C = D$. If now $G' \models b \in \mathcal{G}^C_x \cap \mathcal{G}^D_x$, necessarily $G' \models a \in \mathcal{G}^C_x \cap \mathcal{G}^D_x$, whence $F \models a \in \mathcal{G}^C_x \cap \mathcal{G}^D_x$ and $C = D$.

   To see Requirement 8, $y \in \mathcal{M}^E_x \to \nu(x) \prec_E \nu(y)$, we reason as follows. Again, we first note that $\forall\, x, y{\in}W\ [G' \models y \in \mathcal{M}^C_x$ iff. $F \models y \in \mathcal{M}^C_x]$ holds for any C. We only need to consider the new element, that is, $b \in \mathcal{M}^E_x$. If $x = a$ and $E = B$, we get the property by choice of $\nu(b)$.

For $x \neq a$, we consider two cases. Either $a \in \mathcal{M}_x^E$ or $a \notin \mathcal{M}_x^E$. In the first case, we get by the fact that F is a labeled **IL**M-frame $\nu(x) \prec_E \nu(a)$. But $\nu(a) \prec \nu(b)$, whence $\nu(x) \prec_E \nu(b)$. In the second necessarily for some $a' \in \mathcal{M}_x^E$ we have $a' S^{\mathsf{tr}} a$. But now $\nu(a') \subseteq_\Box \nu(a)$. Clearly $\nu(x) \prec_E \nu(a') \subseteq_\Box \nu(a) \prec \nu(b) \to \nu(x) \prec_E \nu(b)$.

2. With Lemma 6.5 we extend G' to an adequate labeled **IL**M-frame G. It is now obvious that both of the invariants hold on G. The first one holds due to Corollary 6.6. The other is just included in the definition of **IL**M-frames. Obviously, $\langle a, \neg(A \rhd B)\rangle$ is not a problem any more in G.

Deficiencies. Again, any deficiency $\langle a, b, C \rhd D\rangle$ in F will be eliminated in two steps.

1. We first define B to be the formula such that $b \in \mathcal{C}_a^B$. If such a B does not exist, we take B to be $\bot$. Note that if such a B does exist, it must be unique by Property 4 of Definition 5.1. By Lemma 2.8, or just by the fact that F is an **IL**M-frame, we have that $\nu(a) \prec_B \nu(b)$.

   By Lemma 6.7 we can now find a Δ' such that $\nu(a) \prec_B \Delta' \ni D, \Box\neg D$ and $\nu(b) \subseteq_\Box \Delta'$. We fix some $c \notin W$ and define

   $$G' := \langle W, R \cup \{\langle a, c\rangle\}, S \cup \{\langle a, b, c\rangle\}, \nu \cup \{\langle c, \Delta'\rangle\}\rangle.$$

   To see that G' is indeed a quasi-**IL**M-frame, again eight properties should be checked. But all of these are fairly routine.

   For Property 4 it is good to remark that, if $c \in \mathcal{G}_x^A$, then necessarily $b \in \mathcal{G}_x^A$ or $a \in \mathcal{G}_x^A$.

   To see Property 8, we reason as follows. We only need to consider $c \in \mathcal{M}_x^A$. This is possible if $x = a$ and $b \in \mathcal{M}_a^A$, or if for some $y \in \mathcal{M}_x^A$ we have $y S^{\mathsf{tr}} a$, or if $a \in \mathcal{M}_x^A$. In the first case, we get that $b \in \mathcal{M}_a^A$, and thus also $b \in \mathcal{C}_a^A$ as F is an **IL**M-frame. Thus, by Property 4, we see that $A = B$. But Δ' was chosen such that $\nu(a) \prec_B \Delta'$. In the second case we see that $\nu(x) \prec_A \nu(y) \subseteq_\Box \nu(a) \prec \nu(c)$ whence $\nu(x) \prec_A \nu(c)$. In the third case we have $\nu(x) \prec_A \nu(a) \prec \nu(c)$, whence $\nu(x) \prec_A \nu(c)$.

2. Again, G' is closed off under the frame conditions with Lemma 6.5. Clearly, $\langle a, b, C \rhd D\rangle$ is not a deficiency on G.

Rounding up One of our invariants is just the **IL**M frame condition. Clearly this invariant is preserved under taking unions of bounded chains. The closure satisfies the invariants. ⊣

6.3 Admissible rules

With the completeness at hand, a lot of reasoning about **IL**M gets easier. This holds in particular for derived/admissible rules of **IL**M.

Lemma 6.9.

(*i*) $\mathbf{IL}\mathsf{M} \vdash \Box A \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash A$

(*ii*) $\mathbf{IL}\mathsf{M} \vdash \Box A \vee \Box B \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \Box A$ *or* $\mathbf{IL}\mathsf{M} \vdash \Box B$

(*iii*) $\mathbf{IL}\mathsf{M} \vdash A \rhd B \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash A \to B \vee \Diamond B$.

(*iv*) $\mathbf{IL}\mathsf{M} \vdash A \rhd B \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \Diamond A \to \Diamond B$

(*v*) *Let A_i be formulae such that* $\mathbf{IL}\mathsf{M} \nvdash \neg A_i$. *Then*
$\mathbf{IL}\mathsf{M} \vdash \bigwedge \Diamond A_i \to A \rhd B \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash A \rhd B$.

(*vi*) $\mathbf{IL}\mathsf{M} \vdash A \vee \Diamond A \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \Box\bot \to A$

(*vii*) $\mathbf{IL}\mathsf{M} \vdash \top \rhd A \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \Box\bot \to A$

Proof. (*i*). $\mathbf{IL}\mathsf{M} \vdash A \Rightarrow \mathbf{IL}\mathsf{M} \vdash \Box A$ by necessitation. Now suppose $\mathbf{IL}\mathsf{M} \vdash \Box A$. We want to see $\mathbf{IL}\mathsf{M} \vdash A$. Thus, we take an arbitrary model $M = \langle W, R, S, \Vdash \rangle$ and world $m \in M$. If there is an m_0 with $M \models m_0 R m$, then $M, m_0 \Vdash \Box A$, whence $M, m \Vdash A$. If there is no such m_0, we define (we may assume $m_0 \notin W$)

$$\begin{aligned} M' := \quad & \langle W \cup \{m_0\}, R \cup \{\langle m_0, w\rangle \mid w \in W\}, \\ & S \cup \{\langle m_0, x, y\rangle \mid \langle x, y\rangle \in R \text{ or } x{=}y \in W\}, \Vdash\rangle. \end{aligned}$$

Clearly, M' is an $\mathbf{IL}\mathsf{M}$-model too (the $\mathbf{IL}\mathsf{M}$ frame conditions in the new cases follows from the transitivity of R), whence $M', m_0 \Vdash \Box A$ and thus $M', m \Vdash A$. By the construction of M' and by Lemma 2.12 we also get $M, m \Vdash A$.

(*ii*)."$\Leftarrow$" is easy. For the other direction we assume $\mathbf{IL}\mathsf{M} \nvdash \Box A$ and $\mathbf{IL}\mathsf{M} \nvdash \Box B$ and set out to prove $\mathbf{IL}\mathsf{M} \nvdash \Box A \vee \Box B$. By our assumption and by completeness, we find $M_0, m_0 \Vdash \Diamond\neg A$ and $M_1, m_1 \Vdash \Diamond\neg B$. We define (for some $r \notin W_0 \cup W_1$)

$$\begin{aligned} M := \quad & \langle W_0 \cup W_1 \cup \{r\}, R_0 \cup R_1 \cup \{\langle r, x\rangle \mid x \in W_0 \cup W_1\}, \\ & S_0 \cup S_1 \cup \{\langle r, x, y\rangle \mid x{=}y{\in}W_0 \cup W_1 \text{ or } \langle x, y\rangle{\in}R_0 \text{ or } \langle x, y\rangle{\in}R_1\}, \Vdash\rangle. \end{aligned}$$

Now, M is an $\mathbf{IL}\mathsf{M}$-model and $M, r \Vdash \Diamond\neg A \wedge \Diamond\neg B$ as is easily seen by Lemma 2.12. By soundness we get $\mathbf{IL}\mathsf{M} \nvdash \Box A \vee \Box B$.

(*iii*)."$\Leftarrow$" goes as follows. $\vdash A \to B \vee \Diamond B \Rightarrow \vdash \Box(A \to B \vee \Diamond B) \Rightarrow \vdash A \rhd B \vee \Diamond B \Rightarrow \vdash A \rhd B$. For the other direction, suppose that $\nvdash A \to B \vee \Diamond B$. Thus, we can find a model $M = \langle W, R, S, \Vdash\rangle$ and $m \in M$ with $M, m \Vdash A \wedge \neg B \wedge \Box\neg B$. We now define (with $r \notin W$)

$$\begin{aligned} M' := \quad & \langle W \cup \{r\}, R \cup \{\langle r, x\rangle \mid x{=}m \text{ or } \langle m, x\rangle \in R\}, \\ & S \cup \{\langle r, x, y\rangle \mid (x{=}y \text{ and } (\langle m, x\rangle{\in}R \text{ or } x{=}m)) \text{ or } \langle m, x\rangle, \langle x, y\rangle{\in}R\}, \Vdash\rangle. \end{aligned}$$

It is easy to see that M' is an $\mathbf{IL}\mathsf{M}$-model. By Lemma 2.12 we see that $M', x \Vdash \varphi$ iff. $M, x \Vdash \varphi$ for $x \in W$. It is also not hard to see that $M', r \Vdash \neg(A \rhd B)$. For, we have $rRm \Vdash A$. By definition, $mS_r y \to (m{=}y \vee mRy)$ whence $y \nVdash B$.

(iv). By the J4 axiom, we get one direction for free. For the other direction we reason as follows. Suppose $\mathbf{IL}\mathsf{M} \nvdash A \rhd B$. Then we can find a model $M = \langle W, R, S, \Vdash\rangle$ and a world l such that $M, l \Vdash \neg(A \rhd B)$. As $M, l \vdash \neg(A \rhd B)$, w can find some $m \in M$ with $lRm \Vdash A \wedge \neg B \wedge \Box\neg B$. We now define (with $r \notin W$)

$$\begin{array}{rl} M' := & \langle W \cup \{r\}, R \cup \{\langle r, x\rangle \mid x{=}m \text{ or } \langle m, x\rangle \in R\}, \\ & S \cup \{\langle r, x, y\rangle \mid (x{=}y \text{ and } (\langle m, x\rangle{\in}R \text{ or } x{=}m)) \text{ or } \langle m, x\rangle, \langle x, y\rangle{\in}R\}, \Vdash\rangle. \end{array}$$

It is easy to see that M' is an $\mathbf{IL}\mathsf{M}$-model. Lemma 2.12 and general knowledge about $\mathbf{IL}\mathsf{M}$ tells us that the generated submodel from l is a witness to the fact that $\mathbf{IL}\mathsf{M} \nvdash \Diamond A \to \Diamond B$.[13]

(v). The "$\Leftarrow$" direction is easy. For the other direction we reason as follows.[14]

We assume that $\nVdash A \rhd B$ and set out to prove $\nVdash \bigwedge \Diamond A_i \to A \rhd B$. As $\nVdash A \rhd B$, we can find $M, r \Vdash \neg(A \rhd B)$. By Lemma 2.12 we may assume that r is a root of M. For all i, we assumed $\nVdash \neg A_i$, whence we can find rooted models $M_i, r_i \Vdash A_i$. As in the other cases, we define a model $\tilde{M}$ that arises by gluing r under all the r_i. Clearly we now see that $\tilde{M}, r \Vdash \bigwedge \Diamond A_i \wedge \neg(A \rhd B)$.

(vi). First, suppose that $\mathbf{IL}\mathsf{M} \vdash \Box\bot \to A$. Then, from $\mathbf{IL}\mathsf{M} \vdash \Box\bot \vee \Diamond\top$, the observation that $\mathbf{IL}\mathsf{M} \vdash \Diamond\top \leftrightarrow \Diamond\Box\bot$ and our assumption, we get $\mathbf{IL}\mathsf{M} \vdash A \vee \Diamond A$.

For the other direction, we suppose that $\mathbf{IL}\mathsf{M} \nVdash \Box\bot \to A$. Thus, we have a counter model M and some $m \in M$ with $m \Vdash \Box\bot, \neg A$. Clearly, at the submodel generated from m, that is, a single point, we see that $\neg A \wedge \Box\neg A$ holds. Consequently $\mathbf{IL}\mathsf{M}\neg \vdash A \vee \Diamond A$.

(vii). This follows immediately from (vi) and (iii).

$\dashv$

Note that, as $\mathbf{IL}\mathsf{M}$ is conservative over $\mathbf{GL}$, all of the above statements not involving $\rhd$ also hold for $\mathbf{GL}$. The same holds for derived statements. For example, from Lemma 6.9 we can combine (iii) and (iv) to obtain $\mathbf{IL}\mathsf{M} \vdash A \to B \vee \Diamond B \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \Diamond A \to \Diamond B$. Consequently, the same holds true for $\mathbf{GL}$.

6.4 Decidability

It is well known that $\mathbf{IL}\mathsf{M}$ has the finite model property. It is not hard to re-use worlds in the presented construction method so that we would end up with a finite counter model. Actually, this is precisely what has been done in [Joo98]. In that paper, one of the invariants was "there are no deficiencies". We have chosen not to include this invariant in our presentation, as this omission simplifies the presentation. Moreover, for our purposes the completeness without the finite model property obtained via our construction method suffices.

Our purpose to include a new proof of the well known completeness of $\mathbf{IL}\mathsf{M}$ is twofold. On the one hand the new proof serves well to expose the construction

[13] This proof is similar to the proof of (iii). However, it is not the case that one of the two follows easily from the other.

[14] By a similar reasoning we can prove $\vdash \bigwedge \neg(C_i \rhd D_i) \to A \rhd B \Leftrightarrow \vdash A \rhd B$.

method. On the other hand, it is an indispensable ingredient in proving Theorem 7.5.

7 Σ_1-sentences

In this section we will answer the question which modal interpretability sentences are in T provably Σ_1 for any realization. We call these sentences essentially Σ_1-sentences. We shall answer the question only for T an essentially reflexive theory.

This question has been solved for provability logics by Visser in [Vis95]. In [dJP96], de Jongh and Pianigiani gave an alternative solution by using the logic $\mathbf{IL}\mathsf{M}$. Our proof shall use their proof method.

We will perform our argument fully in $\mathbf{IL}\mathsf{M}$. It is very tempting to think that our result would be an immediate corollary from for example [Gor03], [Jap94] or [Ign93]. This would be the case, if a construction method were worked out for the logics from these respective papers. In [Gor03] a sort construction method is indeed worked out. This construction method should however be a bit sharpened to suit our purposes. Moreover that sharpening would essentially reduce to the solution we present here.

7.1 Model construction

Throughout this subsection, unless mentioned otherwise, T will be an essentially reflexive recursively enumerable arithmetical theory. By Theorem 2.26 we thus know that $\mathbf{IL}(\mathrm{T}) = \mathbf{IL}\mathsf{M}$. Let us first say more precisely what we mean by an essentially Σ_1-sentence.

Definition 7.1. A modal sentence φ is called an essentially Σ_1-sentence, if $\forall *\ \varphi^* \in \Sigma_1(T)$. Likewise, a formula φ is essentially Δ_1 if $\forall *\ \varphi^* \in \Delta_1(T)$

If φ is an essentially Σ_1-formula we will also write $\varphi \in \Sigma_1(T)$. Analogously for $\Delta_1(T)$.

Theorem 7.2. *Modulo modal logical equivalence, there exist just two essentially Δ_1-formulas. That is, $\Delta_1(T) = \{\top, \bot\}$.*

Proof. Let φ be a modal formula. If $\varphi \in \Delta_1(T)$, then, by provably Σ_1-completeness, both $\forall *\ T \vdash \delta^* \to \Box\delta^*$ and $\forall *\ T \vdash \neg\delta^* \to \Box\neg\delta^*$. Consequently $\forall *\ T \vdash \Box\delta^* \vee \Box\neg\delta^*$. Thus, $\forall *\ T \vdash (\Box\delta \vee \Box\neg\delta)^*$ whence $\mathbf{IL}\mathsf{M} \vdash \Box\delta \vee \Box\neg\delta$. By Lemma 6.9 we see that $\mathbf{IL}\mathsf{M} \vdash \delta$ or $\mathbf{IL}\mathsf{M} \vdash \neg\delta$. $\dashv$

We proved Theorem 7.2 for the interpretability logic of essentially reflexive theories. It is not hard to see that the theorem also holds for finitely axiomatizable theories. The only ingredients that we need to prove this are [$\mathbf{IL}\mathsf{P} \vdash \Box A \vee \Box B$ iff. $\mathbf{IL}\mathsf{P} \vdash \Box A$ or $\mathbf{IL}\mathsf{P} \vdash \Box B$] and [$\mathbf{IL}\mathsf{P} \vdash \Box A$ iff. $\mathbf{IL}\mathsf{P} \vdash A$]. As these two admissible rules also hold for $\mathbf{GL}$, we see that Theorem 7.2 also holds for $\mathbf{GL}$.

Lemma 7.3. *If $\varphi \in \Sigma_1(T)$, then, for any p and q, we have* $\mathbf{IL}\mathsf{M} \vdash p \rhd q \to p \wedge \varphi \rhd q \wedge \varphi$.

Before we come to prove the main theorem of this section, we first need an additional lemma.

Lemma 7.4. *Let Δ_0 and Δ_1 be maximal* $\mathbf{IL}\mathsf{M}$*-consistent sets. There is a maximal* $\mathbf{IL}\mathsf{M}$*-consistent set Γ such that $\Gamma \prec \Delta_0, \Delta_1$.*

Proof. We show that $\Gamma' := \{\Diamond A \mid A \in \Delta_0\} \cup \{\Diamond B \mid B \in \Delta_1\}$ is consistent. Assume for a contradiction that Γ' were not consistent. Then, by compactness, for finitely many A_i and B_j,

$$\bigwedge_{A_i \in \Delta_0} \Diamond A_i \wedge \bigwedge_{B_j \in \Delta_1} \Diamond B_j \vdash \bot$$

or equivalently

$$\vdash \bigvee_{A_i \in \Delta_0} \Box\neg A_i \vee \bigvee_{B_j \in \Delta_1} \Box\neg B_j.$$

By Lemma 6.9 we see that then either $\vdash \neg A_i$ for some i, or $\vdash \neg B_j$ for some j. This contradicts the consistency of Δ_0 and Δ_1. ⊣

Theorem 7.5. $\varphi \in \Sigma_1(T) \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \varphi \leftrightarrow \bigvee_{i \in I} \Box C_i$ *for some* $\{C_i\}_{i \in I}$.

Proof. Let φ be a formula that is not equivalent to a disjunction of $\Box$-formulas. According to Lemma 7.7 we can find MCS's Δ_0 and Δ_1 with $\varphi \in \Delta_0 \subseteq_\Box \Delta_1 \ni \neg\varphi$. By Lemma 7.4 we find a $\Gamma \prec \Delta_0, \Delta_1$. We define:

$$G := \langle\{m_0, l, r\}, \{\langle m_0, l\rangle, \langle m_0, r\rangle\}, \{\langle m_0, l, r\rangle\}, \{\langle m_0, \Gamma\rangle, \langle l, \Delta_0\rangle, \langle r, \Delta_1\rangle\}\rangle.$$

We will apply a slightly generalized version of the main lemma to this frame quasi-$\mathbf{IL}\mathsf{M}$-frame G. The finite set $\mathcal{D}$ of sentences is the smallest set of sentences that contains φ and that is closed under taking subformulas and single negations. The invariants are the following.

$$\mathcal{I} := \left\{ \begin{array}{l} xRy \wedge x \neq m_0 \to \exists\, A {\in} ((\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}) \\ uS_x vRw \to uRw \end{array} \right.$$

In the proof of Theorem 6.8 we have seen that we can eliminate both problems and deficiencies while conserving the invariants. The main lemma now gives us an $\mathbf{IL}\mathsf{M}$-model M with $M, l \Vdash \varphi$, $M, r \Vdash \neg\varphi$ and $lS_{m_0}r$. We now pick two fresh variables p and q. We define p to be true only at l and q only at r. Clearly $m_0 \Vdash \neg(p \rhd q \to p \wedge \varphi \rhd q \wedge \varphi)$, whence by Lemma 7.3 we get $\varphi \notin \Sigma_1(T)$.
⊣

For finitely axiomatized theories T, our theorem does not hold, as also $A \rhd B$ is T-essentially Σ_1. The following theorem says that in this case, $A \rhd B$ is under any T-realization actually equivalent to a special Σ_1-sentence.

Theorem 7.6. *Let T be a finitely axiomatized theory. For all arithmetical formulae α, β there exists a formula ρ with*

$$T \vdash \alpha \rhd_T \beta \leftrightarrow \Box_T \rho.$$

Proof. The proof is a direct corollary of the so-called FGH-theorem. (See [Vis02] for an exposition of the FGH-theorem.) We take ρ satisfying the following fixed point equation.

$$T \vdash \rho \leftrightarrow ((\alpha \rhd_T \beta) \leq \Box_T \rho)$$

By the proof of the FGH-theorem, we now see that

$$T \vdash ((\alpha \rhd_T \beta) \vee \Box_T \bot) \leftrightarrow \Box_T \rho.$$

But clearly $T \vdash ((\alpha \rhd_T \beta) \vee \Box_T \bot) \leftrightarrow \alpha \rhd_T \beta$. ⊣

7.2 The Σ-lemma

We can say that the proof of Theorem 7.5 contained three main ingredients; Firstly, the main lemma; Secondly the modal completeness theorem for **IL**M via the construction method and; Thirdly the Σ-lemma. In this subsection we will prove the Σ-lemma and remark that it is in a sense optimal.

Lemma 7.7. *If φ is a formula not equivalent to a disjunction of $\Box$-formulas. Then there exist maximal* **IL**X*-consistent sets Δ_0, Δ_1 such that $\varphi \in \Delta_0 \subseteq_\Box \Delta_1 \ni \neg\varphi$.*

Proof. As we shall see, the reasoning below holds not only for **IL**X, but for any extension of **GL**. We define

$$\begin{aligned} \Box_\vee &:= \{ \bigvee_{0 \leq i < n} \Box D_i \mid n \geq 0, \text{each } D_i \text{ an } \mathbf{IL}\mathsf{X}\text{-formula}\}, \\ \Box_{\text{con}} &:= \{Y \subseteq \Box_\vee \mid \{\neg\varphi\} + Y \text{ is consistent and maximally such}\}. \end{aligned}$$

Let us first observe a useful property of the sets Y in $\Box_{\text{con}}$.

$$\bigvee_{i=0}^{n-1} \sigma_i \in Y \Rightarrow \exists\, i{<}n \ \sigma_i \in Y. \tag{1}$$

To see this, let $Y \in \Box_{\text{con}}$ and $\bigvee_{i=0}^{n-1} \sigma_i \in Y$. Then for each $i{<}n$ we have $\sigma_i \in \Box_\vee$ and for some $i{<}n$ we must have σ_i consistent with Y (otherwise $\{\neg\varphi\} + Y$ would prove $\bigwedge_{i=0}^{n-1} \neg\sigma_i$ and be inconsistent). And thus by the maximality of Y we must have that some σ_i is in Y. This establishes (1).

Claim. *For some $Y \in \Box_{\text{con}}$ the set*

$$\{\varphi\} + \{\neg\sigma \mid \sigma \in \Box_\vee - Y\}$$

is consistent.

Proof of the claim. Suppose the claim were false. We will derive a contradiction with the assumption that φ is not equivalent to a disjunction of $\Box$-formulas. If the claim is false, then we can choose for each $Y \in \Box_{\mathrm{con}}$ a finite set $Y^{\mathrm{fin}} \subseteq \Box_{\vee} - Y$ such that

$$\{\varphi\} + \{\neg\sigma \mid \sigma \in Y^{\mathrm{fin}}\} \tag{2}$$

is inconsistent. Thus, certainly for each $Y \in \Box_{\mathrm{con}}$

$$\vdash \varphi \rightarrow \bigvee_{\sigma \in Y^{\mathrm{fin}}} \sigma. \tag{3}$$

Now we will show that:

$$\{\neg\varphi\} + \{\bigvee_{\sigma \in Y^{\mathrm{fin}}} \sigma \mid Y \in \Box_{\mathrm{con}}\} \text{ is inconsistent.} \tag{4}$$

For, suppose (4) were not the case. Then for some $S \in \Box_{\mathrm{con}}$

$$\{\bigvee_{\sigma \in Y^{\mathrm{fin}}} \sigma \mid Y \in \Box_{\mathrm{con}}\} \subseteq S.$$

In particular we have $\bigvee_{\sigma \in S^{\mathrm{fin}}} \sigma \in S$. But for all $\sigma \in S^{\mathrm{fin}}$ we have $\sigma \notin S$. Now by (1) we obtain a contradiction and thus we have shown (4).

So we can select some finite $\Box_{\mathrm{con}}^{\mathrm{fin}} \subseteq \Box_{\mathrm{con}}$ such that

$$\vdash (\bigwedge_{Y \in \Box_{\mathrm{con}}^{\mathrm{fin}}} \bigvee_{\sigma \in Y^{\mathrm{fin}}} \sigma) \rightarrow \varphi. \tag{5}$$

By (3) we also have

$$\vdash \varphi \rightarrow \bigwedge_{Y \in \Box_{\mathrm{con}}^{\mathrm{fin}}} \bigvee_{\sigma \in Y^{\mathrm{fin}}} \sigma. \tag{6}$$

Combining (5) with (6) we get

$$\vdash \varphi \leftrightarrow \bigwedge_{Y \in \Box_{\mathrm{con}}^{\mathrm{fin}}} \bigvee_{\sigma \in Y^{\mathrm{fin}}} \sigma.$$

Bringing the right hand side of this equivalence in disjunctive normal form and distributing the $\Box$ over $\wedge$ we arrive at a contradiction with the assumption on φ. $\dashv$

So, we have for some $Y \in \Box_{\mathrm{con}}$ that both the sets

$$\{\varphi\} + \{\neg\sigma \mid \sigma \in \Box_{\vee} - Y\} \tag{7}$$

$$\{\neg\varphi\} + Y \tag{8}$$

are consistent. The lemma follows by taking Δ_0 and Δ_1 extending (7) and (8) respectively. $\dashv$

We have thus obtained $\varphi \in \Delta_0 \subseteq_{\Box} \Delta_1 \ni \neg\varphi$ for some maximal $\mathbf{IL}\mathsf{X}$-consistent sets Δ_0 and Δ_1. The relation $\subseteq_{\Box}$ between Δ_0 and Δ_1 is actually the best we can get among the relations on MCS's that we consider in this paper. We shall see that $\Delta_0 \prec \Delta_1$ is not possible to get in general.

It is obvious that that $p \wedge \Box p$ is not equivalent to a disjunction of $\Box$-formulas. Clearly $p \wedge \Box p \in \Delta_0 \prec \Delta_1 \ni \neg p \vee \Diamond\neg p$ is impossible. In a sense, this reflects the fact that there exist non trivial self-provers, as was shown by Kent ([Ken73]), Guaspari ([Gua83]) and Beklemishev ([Bek93]). Thus, provable Σ_1-completeness, that is $T \vdash \sigma \to \Box\sigma$ for $\sigma \in \Sigma_1(T)$, can not substitute Lemma 7.3.

7.3 Self provers and Σ_1-sentences

A self prover is a sentence φ that implies its own provability. That is, a sentence for which $\vdash \varphi \to \Box\varphi$, or equivalently, $\vdash \varphi \leftrightarrow \varphi \wedge \Box\varphi$. Self provers have been studied intensively amongst others by Kent ([Ken73]), Guaspari ([Gua83]), de Jongh and Pianigiani ([dJP96]). It is easy to see that any $\Sigma_1(T)$-sentence is indeed a self prover. We shall call such a self prover a *trivial self prover*.

In [Gua83], Guaspari has shown that there are many non-trivial self provers around. The most prominent example is probably $p \wedge \Box p$. But actually, any formula φ will generate a self prover $\varphi \wedge \Box\varphi$, as clearly $\varphi \wedge \Box\varphi \to \Box(\varphi \wedge \Box\varphi)$.

Definition 7.8. A formula φ is called a trivial self prover generator, we shall write t.s.g., if $\varphi \wedge \Box\varphi$ is a trivial self prover. That is, if $\varphi \wedge \Box\varphi \in \Sigma_1(T)$.

Obviously, a trivial self prover is also a t.s.g. But there also exist other t.s.g.'s. The most prominent example is probably $\Box\Box p \to \Box p$. A natural question is to ask for an easy characterization of t.s.g.'s. In this subsection we will give such a characterization for **GL**. In the rest of this subsection, $\vdash$ will stand for derivability in **GL**. We shall often write Σ instead of Σ_1.

We say that a formula ψ is Σ in **GL**, and write $\Sigma(\psi)$, if for any theory T which has **GL** as its provability logic, we have that $\forall * \ \psi^* \in \Sigma_1(T)$.

Theorem 7.9. *We have that $\Sigma(\varphi \wedge \Box\varphi)$ in* **GL** *if and only if the following condition is satisfied.*

For all formulae A_l, φ_l and C_m satisfying 1, 2 and 3 we have that $\vdash \varphi \wedge \Box\varphi \leftrightarrow \bigvee_m \Box C_m$. Here 1-3 are the following conditions.

1. $\vdash \varphi \leftrightarrow \bigvee\!\!\!\bigvee_l (\varphi_l \wedge \Box A_l) \vee \bigvee\!\!\!\bigvee_m \Box C_m$

2. $\nvdash \Box A_l \to \varphi$ *for all* l

3. φ_l *is a non-empty conjunction of literals and* $\Diamond$*-formulas.*

Proof. The $\Leftarrow$ direction is the easiest part. We can always find an equivalent of φ that satisfies 1, 2 and 3. Thus, by assumption, $\varphi \wedge \Box\varphi$ can be written as the disjunction of $\Box$-formulas and hence $\Sigma(\varphi \wedge \Box\varphi)$.

For the $\Rightarrow$ direction we reason as follows. Suppose we can find φ_l, A_l and C_m such that 1, 2 and 3 hold, but

$$\nvdash \varphi \wedge \Box\varphi \leftrightarrow \bigvee\!\!\!\bigvee_m \Box C_m. \quad (*)$$

We can take now $T = \mathrm{PA}$ and reason as follows. As clearly $\vdash \bigvee\!\!\!\bigvee_m \Box C_m \to \varphi \wedge \Box\varphi$, our assumption $(*)$ reduces to $\nvdash \varphi \wedge \Box\varphi \to \bigvee\!\!\!\bigvee_m \Box C_m$. Consequently $\bigvee\!\!\!\bigvee_l (\varphi_l \wedge \Box A_l)$ can not be empty, and for some l and some rooted **GL**-model M, r with root r, we have $M, l \Vdash \Box A_l \wedge \varphi_l$.

We shall now see that $\nvdash \neg\varphi \wedge \Box\varphi \to \Diamond\neg A_l$. For, suppose for a contradiction that

$$\vdash \neg\varphi \wedge \Box\varphi \to \Diamond\neg A_l.$$

Then also $\vdash \Box A_l \to (\Box\varphi \to \varphi)$, whence $\vdash \Box A_l \to \Box(\Box\varphi \to \varphi) \to \Box\varphi$. And by $\Box A_l \to (\Box\varphi \to \varphi)$ again, we get $\vdash \Box A_l \to \varphi$ which contradicts 2. We must conclude that indeed $\nvdash \neg\varphi \wedge \Box\varphi \to \Diamond\neg A_l$, and thus we have a rooted tree model N, r for **GL** with $N, r \Vdash \neg\varphi, \Box\varphi, \Box A_l$.

We can now "glue" a world w below l and r, set lS_wr and consider the smallest **IL**M-model extending this. We have depicted this construction in Figure 1. Let us also give a precise definition. If $M := \langle W_0, R_0, \Vdash_0\rangle$ and $N := \langle W_1, R_1, \Vdash_1\rangle$, then we define

$$\begin{aligned} L := \quad & \langle W_0 \cup W_1, R_0 \cup R_1 \cup \{\langle w, x\rangle \mid x \in W_0 \cup W_1\} \cup \{\langle l, y\rangle \mid N \models rRy\}, \\ & \{\langle w, l, r\rangle\} \cup \{\langle x, y, z\rangle \mid L \models xRyR^*z\}, \Vdash_0 \cup \Vdash_1\rangle. \end{aligned}$$

We observe that, by Lemma 2.12 $L, r \Vdash \Box\varphi \wedge \Box A_l \wedge \neg\varphi$ and $L \models rRx \Rightarrow L, x \Vdash \varphi \wedge A_l$. Also, if $L \models lRx$, then $L, x \Vdash \varphi \wedge A_i$, whence $L, l \Vdash \Box\varphi \wedge \Box A_l$. As $M, l \Vdash \varphi_l$ and φ_l only contains literals and and diamond-formulas, we see that $L, l \Vdash \varphi_l$, whence $L, l \Vdash \varphi \wedge \Box\varphi$. As $L, r \Vdash \neg\varphi \wedge \Box\varphi$ we see that $L, w \Vdash \neg\Sigma(\varphi \wedge \Box\varphi)$.

As in the proof of Theorem 7.5, we can take some fresh p and q and define p to hold only at l and q to hold only at r. Now, clearly $w \nVdash p \rhd q \to p \wedge (\varphi \wedge \Box\varphi) \rhd q \wedge (\varphi \wedge \Box\varphi)$, whence, by Lemma 7.3 we conclude $\neg\Sigma(\varphi \wedge \Box\varphi)$. $\dashv$

To conclude this subsection, we remain in **GL** and shall settle the question for which φ we have that

$$\Sigma(\varphi \wedge \Box\varphi) \ \& \ \Sigma(\varphi \wedge \Box\neg\varphi) \Rightarrow \Sigma(\varphi). \quad (\dagger)$$

We shall see how this question can be reduced to the characterization of t.s.g.'s.

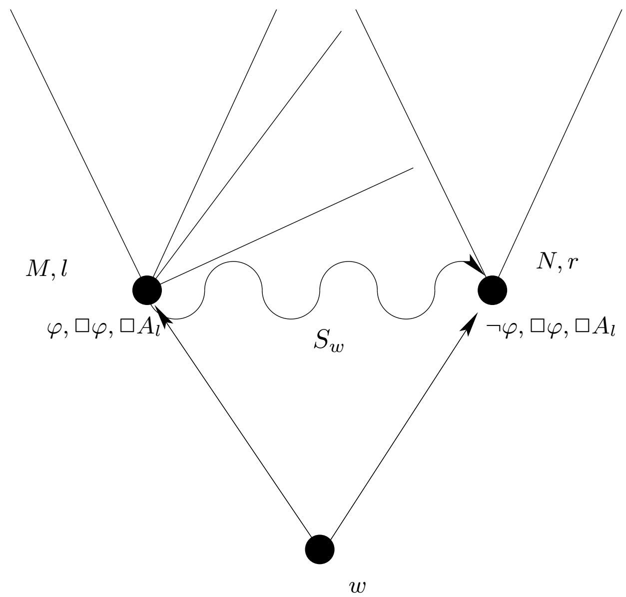

Figure 1: T.s.g.'s

Lemma 7.10.

For some (possibly empty) $\bigvee\!\!\!\bigvee_i \Box C_i$ *we have* $\vdash \varphi \wedge \Box\neg\varphi \leftrightarrow \bigvee\!\!\!\bigvee_i \Box C_i$
iff.
$\vdash \Box\bot \rightarrow \varphi$ *or* $\vdash \neg\varphi$

Proof. For non-empty $\bigvee\!\!\!\bigvee_i \Box C_i$ we have the following.

$$\begin{array}{ll}
\vdash \varphi \wedge \Box\neg\varphi \leftrightarrow \bigvee\!\!\!\bigvee_i \Box C_i & \Rightarrow \\
\vdash \Diamond(\varphi \wedge \Box\neg\varphi) \leftrightarrow \Diamond(\bigvee\!\!\!\bigvee_i \Box C_i) & \Rightarrow \\
\vdash \Diamond\varphi \leftrightarrow \Diamond\top & \Rightarrow \\
\vdash \Box\bot \rightarrow \varphi &
\end{array}$$

Here, the final step in the proof comes from Lemma 6.9.

On the other hand, if $\vdash \Box\bot \rightarrow \varphi$, we see that $\vdash \neg\varphi \rightarrow \Diamond\top$ and thus $\Box\neg\varphi \rightarrow \Box\bot$, whence $\vdash \varphi \wedge \Box\neg\varphi \leftrightarrow \Box\bot$.

In case of the empty disjunction we get $\vdash \varphi \wedge \Box\neg\varphi \leftrightarrow \bot$. Then also $\vdash \Box\neg\varphi \rightarrow \neg\varphi$ and by Löb $\vdash \neg\varphi$. And conversely, if $\vdash \neg\varphi$, then $\vdash \varphi \wedge \Box\neg\varphi \leftrightarrow \bot$, and $\bot$ is just the empty disjunction.

The proof actually gives some additional information. If $\Sigma(\varphi \wedge \Box\neg\varphi)$ then either ($\vdash \neg\varphi$ and $\vdash (\varphi \wedge \Box\neg\varphi) \leftrightarrow \bot$), or ($\vdash \Box\bot \rightarrow \varphi$ and $\vdash (\varphi \wedge \Box\neg\varphi) \leftrightarrow \Box\bot$). $\dashv$

Lemma 7.11.

$\Sigma(\varphi \wedge \Box\varphi) \wedge \Sigma(\varphi \wedge \Box\neg\varphi) \Rightarrow \Sigma(\varphi)$
iff.
$\Sigma(\varphi \wedge \Box\varphi) \Rightarrow \Sigma(\varphi)$ *or* $\vdash \varphi \rightarrow \Diamond\top$

Proof. $\Uparrow$. Clearly, if $\Sigma(\varphi \wedge \Box\varphi) \Rightarrow \Sigma(\varphi)$, also $\Sigma(\varphi \wedge \Box\varphi) \wedge \Sigma(\varphi \wedge \Box\neg\varphi) \Rightarrow \Sigma(\varphi)$. Thus, suppose $\vdash \varphi \rightarrow \Diamond\top$, or put differently $\vdash \Box\bot \rightarrow \neg\varphi$. If now $\vdash \neg\varphi$, then clearly $\Sigma(\varphi)$, whence $\Sigma(\varphi \wedge \Box\varphi) \wedge \Sigma(\varphi \wedge \Box\neg\varphi) \Rightarrow \Sigma(\varphi)$, so, we may assume that $\nvdash \neg\varphi$. It is clear that now $\neg\Sigma(\varphi \wedge \Box\neg\varphi)$. For, suppose $\Sigma(\varphi \wedge \Box\neg\varphi)$, then by Lemma 7.10 we see $\vdash \Box\bot \rightarrow \varphi$, whence $\vdash \Diamond\top$. Quod non. Thus, $\vdash \Box\bot \rightarrow \neg\varphi \Rightarrow \neg\Sigma(\varphi \wedge \Box\neg\varphi)$ and thus certainly $\Sigma(\varphi \wedge \Box\varphi) \wedge \Sigma(\varphi \wedge \Box\neg\varphi) \Rightarrow \Sigma(\varphi)$.

$\Downarrow$. Suppose $\Sigma(\varphi \wedge \Box\varphi) \wedge \neg\Sigma(\varphi)$ and $\nvdash \Box\bot \rightarrow \neg\varphi$. To obtain our result, we only have to prove $\Sigma(\varphi \wedge \Box\neg\varphi)$.

As $\nvdash \Box\bot \rightarrow \neg\varphi$, also $\nvdash \neg\varphi \vee \Diamond\neg\varphi$. Thus, under the assumption that $\Sigma(\varphi \wedge \Box\varphi)$, we can find (a non-empty collection of) C_i with $\vdash \varphi \wedge \Box\varphi \leftrightarrow \bigvee_i \Box C_i$. In this case, clearly $\vdash \Box\bot \rightarrow \bigvee_i \Box C_i \rightarrow \varphi$, whence, by Lemma 7.10 we conclude $\Sigma(\varphi \wedge \Box\neg\varphi)$. $\dashv$

8 The logic $\mathbf{IL}\mathsf{M}_0$

This section is devoted to showing the following theorem.[15]

Theorem 8.1. $\mathbf{IL}\mathsf{M}_0$ *is a complete logic.*

In the light of Remark 4.19 a proof of Theorem 8.1 boils down to giving the four ingredients mentioned there. Sections 8.3, 8.4, 8.5, 8.6 and 8.7 below contain those ingredients. Before these main sections, we have in Section 8.2 some preliminaries. We start in Section 8.1 with an overview of the difficulties we encounter during the application of the construction method to $\mathbf{IL}\mathsf{M}_0$.

8.1 Overview of difficulties

In the construction method we repeatedly eliminate problems and deficiencies by extensions that satisfy all the invariants. During these operations we need to keep track of two things.

1. If x has been added to solve a problem in w, say $\neg(A \rhd B) \in \nu(w)$. Then for all y such that xS_wy we have $\nu(w) \prec_B \nu(y)$.
2. If wRx then $\nu(w) \prec \nu(x)$

Item 1. does not impose any direct difficulties. But some do emerge when we try to deal with the difficulties concerning Item 2. So let us see why it is difficult to ensure 2. Suppose we have $wRxRyS_wy'Rz$. The M_0–frame condition (Theorem 8.19) requires that we also have xRz. So, from 2. and the M_0–frame condition we obtain $wRxRyS_wy'Rz \rightarrow \nu(x) \prec \nu(z)$. A sufficient (and in certain sense necessary) condition is,

$$wRxRyS_wy' \rightarrow \nu(x) \subseteq_\Box \nu(y').$$

[15] A proof sketch of this theorem was first given in [Joo98].

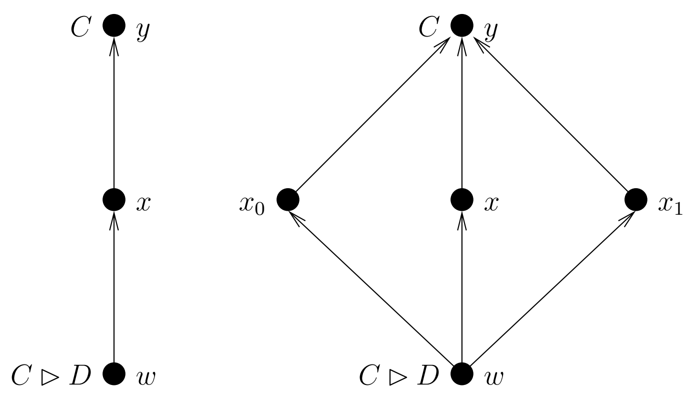

Figure 2: A deficiency in w w.r.t. y

Let us illustrate some difficulties concerning this condition by some examples. Consider the left model in Figure 2. That is, we have a deficiency in w w.r.t. y. Namely, $C \rhd D \in \nu(w)$ and $C \in \nu(y)$. If we solve this deficiency by adding a world y', we thus require that for all x such that $wRxRy$ we have $\nu(x) \subseteq_\Box \nu(y')$. This difficulty is partially handled by Lemma 8.2 below. We omit a proof, but it can easily be given by replacing in the corresponding lemma for **IL**M, applications of the M-axiom by applications of the M_0-axiom.

Lemma 8.2. *Let* Γ, Δ *be* MCS*'s such that* $C \rhd D \in \Gamma$*,* $\Gamma \prec_A \Delta$ *and* $\Diamond C \in \Delta$*. Then there exists some* Δ' *with* $\Gamma \prec_A \Delta'$*,* $\Box\neg D, D \in \Delta'$ *and* $\Delta \subseteq_\Box \Delta'$*.*

Let us now consider the right most model in Figure 2. We have at least for two different worlds x, say x_0 and x_1, that $wRxRy$. Lemma 8.2 is applicable to $\nu(x_0)$ and $\nu(x_1)$ separately but not simultaneously. In other words we find y'_0 and y'_1 such that $\nu(x_0) \subseteq_\Box \nu(y'_0)$ and $\nu(x_1) \subseteq_\Box \nu(y'_1)$. But we actually want one single y' such that $\nu(x_0) \subseteq_\Box \nu(y')$ and $\nu(x_1) \subseteq_\Box \nu(y')$. We shall handle this difficulty by ensuring that it is enough to consider only one of the worlds in between w and y. To be precise, we shall ensure $\nu(x') \subseteq_\Box \nu(x)$ or $\nu(x) \subseteq_\Box \nu(x')$.

But now some difficulties concerning Item 1. occur. In the situations in Figure 2 we were asked to solve a deficiency in w w.r.t. y. As usual, if $w \prec_A y$ then we should be ably to choose a solution y' such that $w \prec_A y'$. But Lemma 8.2 takes only criticallity of x w.r.t. w into account. This issue is solved by ensuring that $wRxRy \in \mathcal{C}^A_w$ implies $\nu(w) \prec_A \nu(x)$.

We are not there yet. Consider the leftmost model in Figure 3. That is, we have a deficiency in w w.r.t. y'. Namely, $C \rhd D \in \nu(w)$ and $C \in \nu(y')$. If we add a world y'' to solve this deficiency, as in the middle model, then by transitivity of S_w we have yS_wy'', as shown in the rightmost model. So, we require that $\nu(x) \subseteq_\Box \nu(y'')$. But we might very well have $\Diamond C \notin \nu(x)$. So the Lemma 8.2 is not applicable.

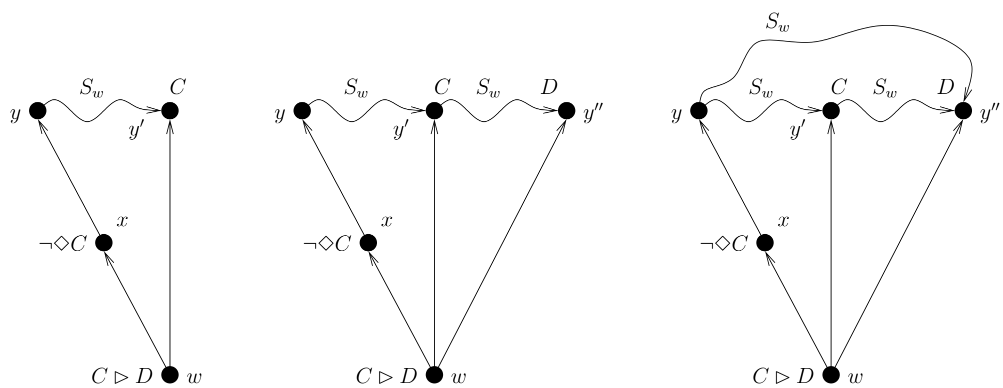

Figure 3: A deficiency in w w.r.t. y'

In Lemma 8.16 we formulate and prove a more complicated version of the Lemma 8.2 which basically says that if we have chosen $\nu(y')$ appropriately, then we can choose $\nu(y'')$ such that $\nu(x) \subseteq_{\Box} \nu(y'')$. And moreover, Lemma 8.16 ensures us that we can, indeed, choose $\nu(y')$ appropriate.

8.2 Preliminaries

Definition 8.3 ($T^{\mathbf{tr}}$, T^*, $T;T'$, T^1, $T^{\geq 2}$, $T \cup T'$). Let T and T' be binary relations on a set W. We fix the following fairly standard notations. T^{tr} is the transitive closure of T; T^* is the transitive reflexive closure of T; $xT;T'y \Leftrightarrow \exists t\; xTtT'y$; $xT^1y \Leftrightarrow xTy \wedge \neg\exists t\; xTtTy$; $xT^{\geq 2}y \Leftrightarrow xTy \wedge \neg(xT^1y)$ and $xT \cup T'y \Leftrightarrow xTy \vee xT'y$.

Definition 8.4 ($\mathcal{S}_w$). Let $F = \langle W, R, S, \nu \rangle$ be a quasi–frame. For each $w \in W$ we define the relation $\mathcal{S}_w$, of pure S_w transitions, as follows.

$$x\mathcal{S}_w y \Leftrightarrow xS_w y \wedge \neg(x = y) \wedge \neg(x(S_w \cup R)^*; R; (S_w \cup R)^* y)$$

Definition 8.5 (Adequate ILM$_0$–frame). Let $F = \langle W, R, S, \nu \rangle$ be an adequate frame. We say that F is an adequate **IL**M$_0$–frame iff. the following additional properties hold.[16]

4. $wRxRyS_w y'Rz \to xRz$
5. $wRxRyS_w y' \to \nu(x) \subseteq_{\Box} \nu(y')$
6. $xS_w y \to x(\mathcal{S}_w \cup R)^* y$
7. $xRy \to x(R^1)^{\mathrm{tr}} y$

[16]One might think that 6. is superfluous. In finite frame this is indeed the case, but in the general case we need it as an requirement.

As usual, when we speak of $\mathbf{IL}\mathsf{M}_0$–frames we shall actually mean an adequate $\mathbf{IL}\mathsf{M}_0$–frame. Below we will construct $\mathbf{IL}\mathsf{M}_0$–frames out of frames belonging to a certain subclass of the class of quasi–frames. (Namely the quasi–$\mathbf{IL}\mathsf{M}_0$–frames, see Definition 8.10 below.) We would like to predict on forehand which extra R relations will be added during this construction. The following definition does just that.

Definition 8.6 ($K(F)$, K). Let $F = \langle W, R, S, \nu\rangle$ be a quasi–frame. We define $K = K(F)$ to be the smallest binary relation on W such that

1. $R \subseteq K$,
2. $K = K^{\mathrm{tr}}$,
3. $wKxK^1y(\mathcal{S}_w)^{\mathrm{tr}}y'K^1z \to xKz$.

Note that for $\mathbf{IL}\mathsf{M}_0$–frames we have $K = R$. The following lemma shows that K satisfies some stability conditions. The lemma will mainly be used to show that whenever we extend R within K, then K does not change.

Lemma 8.7. *Let $F_0 = \langle W, R_0, S, \nu\rangle$ and $F_1 = \langle W, R_1, S, \nu\rangle$ be quasi–frames. If $R_1 \subseteq K(F_0)$ and $R_0 \subseteq K(F_1)$. Then $K(F_0) = K(F_1)$.*

In a great deal of situations we have a particular interest in K^1. To determine some of its properties the following lemma comes in handy. It basically shows that we can compute K by first closing of under the M_0–condition and then take the transitive closure.

Lemma 8.8 (Calculation of K). *Let $F = \langle W, R, S, \nu\rangle$ be a quasi–frame. Let $K = K(F)$ and suppose K conversely well–founded. Let T be a binary relation on W such that*

1. $R \subseteq T^{tr} \subseteq K$,
2. $wT^{tr}xT^1y(\mathcal{S}_w)^{tr}y'T^1z \to xT^{tr}z$.

Then we have the following.

(a) $K = T^{tr}$

(b) $xK^1y \to xTy$

Proof. To see (a), it is enough to see that T^{tr} satisfies the three properties of the definition of K (Definition 8.6). Item (b) follows from (a). $\dashv$

Another entity that changes during the construction of an $\mathbf{IL}\mathsf{M}_0$–frame out of a quasi–frame is the critical cone In accordance with the above definition of $K(F)$, we also like to predict what eventually becomes the critical cone.

Definition 8.9 ($\mathcal{N}^C_w$). For any quasi–frame F we define $\mathcal{N}^C_w$ to be the smallest set such that

1. $\nu(w,x) = C \Rightarrow x \in \mathcal{N}_w^C$,
2. $x \in \mathcal{N}_w^C \wedge x(K \cup S_w)y \Rightarrow y \in \mathcal{N}_w^C$.

In accordance with the notion of a quasi–frame we introduce the notion of a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. This gives sufficient conditions for a quasi–frame to be closeable, not only under the **IL**–frameconditions, but under all the $\mathbf{IL}\mathsf{M}_0$–frameconditions.

Definition 8.10 (Quasi–$\mathbf{IL}\mathsf{M}_0$–frame). A quasi–$\mathbf{IL}\mathsf{M}_0$–frame is a quasi–frame that satisfies the following additional properties.

6. K is conversely well–founded.
7. $xKy \to \nu(x) \prec \nu(y)$
8. $x \in \mathcal{N}_w^A \to \nu(w) \prec_A \nu(x)$
9. $wKxKy(S_w \cup K)^* y' \to \nu(x) \subseteq_\Box \nu(y')$
10. $xS_w y \to x(\mathcal{S}_w \cup R)^* y$
11. $wKxK^1 y(\mathcal{S}_w)^{\mathrm{tr}} y' K^1 z \to x(K^1)^{\mathrm{tr}} z$
12. $xRy \to x(R^1)^{\mathrm{tr}} y$

Lemma 8.11. *If F is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame, then $K = (K^1)^{tr}$.*

Proof. Using Lemma 8.8. ⊣

Lemma 8.12. *Suppose that F is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. Let $K = K(F)$. Let K', K'' and K''' the smallest binary relations on W satifying 1. and 2. of 8.6 and additionaly we have the following.*

3′. $wK'xK'^1 y(\mathcal{S}_w \cup K')^* y' K'^1 z \to xK'z$

3″. $wK''xK''y(\mathcal{S}_w)^{tr} y' K'' z \to xK''z$

3‴. $wK'''xK'''y(S_w \cup K''')^* y' K''' z \to xK'''z$

Then $K = K' = K'' = K'''$.

Proof. Using Lemma 8.11. ⊣

Before we move on, let us first sum up a few comments.

Corollary. *If $F = \langle W, R, S, \nu\rangle$ is an adequate $\mathbf{IL}\mathsf{M}_0$–frame. Then we have the following.*

1. $K(F) = R$

2. $F \models x \in \mathcal{N}_w^A \Leftrightarrow F \models x \in \mathcal{C}_w^A$

3. F is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame

Lemma 8.13 ($\mathbf{IL}\mathsf{M}_0$–closure). *Any quasi–$\mathbf{IL}\mathsf{M}_0$–frame can be extended to an adequate $\mathbf{IL}\mathsf{M}_0$–frame.*

Proof. Given a quasi–$\mathbf{IL}\mathsf{M}_0$–frame F we construct a sequence

$$F = F_0 \subseteq F_1 \subseteq \cdots$$

very similar to the sequence constructed for the **IL** closure of a quasi–frame (Lemma 5.2). The only difference is that we add a fifth entry to the list of imperfections.

(v) $\gamma = \langle 4, w, a, b, b', c\rangle$ with $F_n \models wRaRbS_w b'Rc$ but $F_n \not\models aRc$

In this case we set, of course, $F_{n+1} := \langle W_n, R_n \cup \langle a, c\rangle, S_n, \nu_n\rangle$. First we will show by induction that each F_n is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. Then we show that the union $\hat{F} = \bigcup_{n\geq 0} F_n$, is quasi and satisfies all the $\mathbf{IL}\mathsf{M}_0$ frame conditions.

We assume that F_n is a quasi-$\mathbf{IL}\mathsf{M}_0$-frame and define $K^n := K(F_n)$, $R^n := R^{F_n}$ and $S^n := S^{F_n}$. Quasi-ness of F_{n+1} will follow from Claim 8.13a, and from Claim 8.13b we may conlude that F_{n+1} is indeed a quasi-$\mathbf{IL}\mathsf{M}_0$-frame.

Claim 8.13a. For all w, x, y and A we have the following.

(a) $R^{n+1} \subseteq K^n$

(b) $x(S_w^{n+1} \cup R^{n+1})^* y \Rightarrow x(S_w^n \cup K^n)^* y$

(c) $F_{n+1} \models x \in \mathcal{C}_w^A \Rightarrow F_n \models x \in \mathcal{N}_w^A$.

Proof. We distinguish cases according to which imperfection is dealt with in the step from F_n to F_{n+1}. The only interesting case is the new imperfection which is dealt with by Lemma 8.12, Item 3″. ⊣

Claim 8.13b. For all w, x and A we have the following.

1. $K^{n+1} \subseteq K^n$.
2. $x(S_w^{n+1} \cup K^{n+1})^* y \Rightarrow x(S_w^n \cup K^n)^* y$
3. $F_{n+1} \models x \in \mathcal{N}_w^A \Rightarrow F_n \models x \in \mathcal{N}_w^A$.

Proof. Item 1. follows by Claim 8.13a and Lemma 8.7. Item 2. follows from Item 1. and Claim 8.13a-(b). Item 3. is an immediate corollary of item 2. ⊣

Again, it is not hard to see that $\hat{F} = \bigcup_{n\geq 0} F_n$ is an adequate $\mathbf{IL}\mathsf{M}_0$-frame. ⊣

Lemma 8.14. *Let $F = \langle W, R, S, \nu\rangle$ be a quasi–$\mathbf{IL}\mathsf{M}_0$–frame and $K = K(F)$. Then*

$$xKy \rightarrow \exists z\,(\nu(x) \subseteq_\Box \nu(z) \wedge x(R \cup S)^* zRy).$$

Proof. We define $T := \{(x,y) \mid \exists z\,(\nu(x) \subseteq_\Box \nu(z) \wedge x(R\cup S)^* zRy)\}$. It is not hard to see that T is transitive and that $\{(x,y) \mid \exists t\,(\nu(x) \subseteq_\Box \nu(t) \wedge xT;(S\cup K)^* tTy)\} \subseteq T$. We now define $K' = K \cap T$. We have to show that $K' = K$. As $K' \subseteq K$ is trivial, we will show $K \subseteq K'$.

It is easy to see that K' satisfies properties 1., 2. and 3. of Definition 8.6; It follows on the two observations on T we just made. Since K is the smallest binary relation that satisfies these properties we conclude $K \subseteq K'$. $\dashv$

The next lemma shows that K is a rather stable relation. We show that if we extend a frame G to a frame F such that from worlds in $F - G$ we cannot reach worlds in G, then K on G does not change.

Lemma 8.15. *Let $F = \langle W, R, S, \nu\rangle$ be a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. And let $G = \langle W^-, R^-, S^-, \nu^-\rangle$ be a subframe of F (which means $W^- \subseteq W$, $R^- \subseteq R$, $S^- \subseteq S$ and $\nu^- \subseteq \nu$). If*

(a) for each $f \in W - W^-$ and $g \in W^-$ not $f(R\cup S)g$ and

(b) $R{\restriction}_{W^-} \subseteq K(G)$.

Then $K(G) = K(F){\restriction}_{W^-}$.

Proof. Clearly $K(F){\restriction}_{W^-}$ satisfies the properties 1., 2. and 3. of the definition of $K(G)$ (Definition 8.6). Thus, since K_G is the smallest such relation, we get that $K(G) \subseteq K(F){\restriction}_{W^-}$.

Let $K' = K(F) - (K(F){\restriction}_{W^-} - K(G))$. Using Lemma 8.14 one can show that $K(F) \subseteq K'$. From this it immediately follows that $K(F){\restriction}_{W^-} \subseteq K(G)$. $\dashv$

We finish the basic preliminaries with a somewhat complicated variation of Lemma 4.18.

Lemma 8.16. *Let Γ and Δ be* MCS*'s. $\Gamma \prec_C \Delta$.*

$$P \rhd Q, S_1 \rhd T_1, \ldots, S_n \rhd T_n \in \Gamma \quad \text{and} \quad \Diamond P \in \Delta.$$

There exist $k \le n$. MCS*'s $\Delta_0, \Delta_1, \ldots, \Delta_k$ such that*

- *Each Δ_i lies C-critical above Γ,*
- *Each Δ_i lies $\subseteq_\Box$ above Δ (i.e. $\Delta \subseteq_\Box \Delta_i$),*
- *$Q \in \Delta_0$,*
- *For all $1 \le j \le n$, $S_j \in \Delta_h \Rightarrow$ for some $i \le k$, $T_j \in \Delta_i$.*

Proof. First a definition. For each $I \subseteq \{1, \ldots, n\}$ put

$$\overline{S}_I :\Leftrightarrow \bigwedge\{\neg S_i \mid i \in I\}.$$

The lemma can now be formulated as follows. There exists $I \subseteq \{1, \ldots, n\}$ such that

$$\{Q, \overline{S}_I\} \cup \{\neg B, \Box\neg B \mid B \rhd C \in \Gamma\} \cup \{\Box A \mid \Box A \in \Delta\} \nvdash \bot$$

and, for all $i \notin I$,

$$\{T_i, \overline{S}_I\} \cup \{\neg B, \Box\neg B \mid B \rhd C \in \Gamma\} \cup \{\Box A \mid \Box A \in \Delta\} \nvdash \bot.$$

So let us assume, for a contradiction, that this is false. Then there exist finite sets $\mathcal{A} \subseteq \{A \mid \Box A \in \Delta\}$ and $\mathcal{B} \subseteq \{B \mid B \rhd C \in \Gamma\}$ such that, if we put

$$A :\Leftrightarrow \bigwedge \mathcal{A}, \text{ and } B :\Leftrightarrow \bigvee \mathcal{B},$$

then, for all $I \subseteq \{1, \ldots, n\}$,

$$Q, \overline{S}_I, \Box A, \neg B \wedge \Box\neg B \vdash \bot \tag{9}$$

or,

$$\text{for some } i \notin I, \quad T_i, \overline{S}_I, \Box A, \neg B \wedge \Box\neg B \vdash \bot. \tag{10}$$

We are going to define a permutation $i_1, \ldots, i_n$ of $\{1, \ldots, n\}$ such that if we put $I_k = \{i_j \mid j < k\}$ then

$$T_{i_k}, \overline{S}_{I_k}, \Box A, \neg B \wedge \Box\neg B \vdash \bot. \tag{11}$$

Additionally, we will verify that for each k

$$(9) \text{ does not hold with } I_k \text{ for } I.$$

We will define i_k with induction on k. We define $I_1 = \emptyset$. And by Lemma 4.18, (9) does not hold with $I = \emptyset$. Moreover, because of this, (10) must be true with $I = \emptyset$. So, there exists some $i \in \{1, \ldots, n\}$ such that

$$T_i, \Box A, \neg B \wedge \Box\neg B \vdash \bot.$$

It is thus sufficient to take for i_1, for example, the least such i.

Now suppose i_k has been defined. We will first show that

$$Q, \overline{S}_{I_{k+1}}, \Box A, \neg B \wedge \Box\neg B \nvdash \bot. \tag{12}$$

Let us suppose that this is not so. Then

$$\vdash \Box(Q \to \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_k}). \tag{13}$$

So,

$$\begin{aligned}
\Gamma \vdash P &\rhd Q \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee S_{i_k} && \text{by (13)} \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee T_{i_k} \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee (T_{i_k} \wedge \Box A \wedge \neg B \wedge \Box\neg B \wedge \overline{S}_{I_k}) \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} && \text{by (11)} \\
&\quad \vdots \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee T_{i_1} \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee (T_{i_1} \wedge \Box A \wedge \neg B \wedge \Box\neg B) \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B. && \text{by (11), with } k = 1.
\end{aligned}$$

So by $\mathsf{M_0}$,

$$\Diamond P \wedge \Box A \rhd (\Diamond\neg A \vee B \vee \Diamond B) \wedge \Box A \in \Gamma.$$

But $\Diamond P \wedge \Box A \in \Delta$. So, by Lemma 4.18 there exists some MCS Δ with $\Gamma \prec_C \Delta$ that contains $B \vee \Diamond B$. This is a contradiction, so we have shown (12).

But now, since (12) is indeed true, and thus (9) with I_{k+1} for I is false, (10) must hold. Thus there must exist some $i \notin I_{k+1}$ such that

$$T_i, \overline{S}_{I_{k+1}}, \Box A, \neg B \wedge \Box\neg B \vdash \bot.$$

So we can take for i_{k+1}, for example, the smallest such i.

It is clear that for $I = \{1, 2, \ldots, n\}$, (10) cannot be true. Thus, for $I = \{1, 2, \ldots, n\}$, (9) must be true. This implies

$$\vdash \Box(Q \to \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_n}).$$

Now exactly as above we can show $\Gamma \vdash P \rhd \Diamond\neg A \vee B \vee \Diamond B$. And again as above, this leads to a contradiction. $\dashv$

In order to formulate the invariants needed in the main lemma applied to $\mathbf{IL}\mathsf{M_0}$, we need one more definition and a corollary.

Definition 8.17 ($\subset_1$, $\subset$). Let $F = \langle W, R, S, \nu \rangle$ be a quasi–frame. Let $K = K(F)$. We define $\subset_1$ and $\subset$ as follows.

1. $x \subset_1 y \Leftrightarrow \exists w y' wKxK^1y'(\mathcal{S}_w)^{\mathrm{tr}}y$
2. $x \subset y \Leftrightarrow x(\subset_1 \cup K)^* y$

Corollary 8.18. *Let $F = \langle W, R, S, \nu \rangle$ be a quasi–frame. And let $K = K(F)$.*

1. $x \subset y \wedge yKz \to xKz$
2. *If F is a quasi–$\mathbf{IL}\mathsf{M_0}$–frame, then $x \subset y \Rightarrow \nu(x) \subseteq_\Box \nu(y)$.*

8.3 Frame condition

The following theorem is well known.

Theorem 8.19. *For an $\mathbf{IL}$-frame $F = \langle W, R, S, \nu \rangle$ we have*

$$\forall wxyy'z\ (wRxRyS_wy'Rz \to xRz) \Leftrightarrow F \models \mathsf{M_0}.$$

8.4 Invariants

Let $\mathcal{D}$ be some finite set of formulas, closed under subformulas and single negation.

During the construction we will keep track of the following main–invariants.

$\mathcal{I}_\Box$ for all y, $\{\nu(x) \mid xK^1y\}$ is linearly ordered by $\subseteq_\Box$

$\mathcal{I}_{\mathrm{d}}$ $wK^1x \wedge wK^{\geq 2}x'(S_w \cup K)^*x \to$ 'there does not exists a deficiency in w w.r.t. x'

$\mathcal{I}_S$ $wKxKy(S_w \cup K)^*y' \to$
'the $\subseteq_\Box$-max of $\{\nu(t) \mid wKtK^1y'\}$, if it exists, is $\subseteq_\Box$-larger than $\nu(x)$'

$\mathcal{I}_{\mathcal{N}}$ $wKxKy \wedge y \in \mathcal{N}_w^A \to x \in \mathcal{N}_w^A$

$\mathcal{I}_{\mathcal{D}}$ $xRy \to \exists\, A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}$

$\mathcal{I}_{\mathsf{M}_0}$ All conditions for an adequate $\mathbf{IL}\mathsf{M}_0$–frame hold

In order to ensure that the main–invariants are preserved during the construction we need to consider the following sub–invariants.[17]

$\mathcal{J}_{\mathrm{u}}$ $wK^{\geq 2}x(\mathcal{S}_w)^{\mathrm{tr}}y \wedge wK^{\geq 2}x'(\mathcal{S}_w)^{\mathrm{tr}}y \to x = x'$

$\mathcal{J}_{K^1}$ $wKxK^1y(\mathcal{S}_w)^{\mathrm{tr}}y'K^1z \to xK^1z$

$\mathcal{J}_{\subset}$ $y \subset x \wedge x \subset y \to y = x$

$\mathcal{J}_{\mathcal{N}_1}$ $x(\mathcal{S}_v)^{\mathrm{tr}}y \wedge wKy \wedge x \in \mathcal{N}_w^A \to y \in \mathcal{N}_w^A$

$\mathcal{J}_{\mathcal{N}_2}$ $x(\mathcal{S}_w)^{\mathrm{tr}}y \wedge y \in \mathcal{N}_w^A \to x \in \mathcal{N}_w^A$

$\mathcal{J}_{\nu_1}$ '$\nu(w,y)$ is defined' $\wedge\, vKy \to v \subset w$

$\mathcal{J}_{\nu_2}$ '$\nu(w,y)$ is defined' $\to wK^1y$

$\mathcal{J}_{\nu_4}$ If $x(\mathcal{S}_w)^{tr}y$, then $\nu(w,y)$ is defined

$\mathcal{J}_{\nu_3}$ If $\nu(v,y)$ and $\nu(w,y)$ are defined then $w = v$

What can we say about these invariants? $\mathcal{I}_\Box, \mathcal{I}_S, \mathcal{I}_{\mathcal{N}}$ and $\mathcal{I}_{\mathrm{d}}$ were discussed in Section 8.1. $\mathcal{I}_{\mathsf{M}_0}$ is there to ensure that our final frame is an $\mathbf{IL}\mathsf{M}_0$–frame. About the sub–invariants there is not much to say. They are merely technicalities that ensure that the main–invariants are invariant.

Let us first show that if we have a quasi–$\mathbf{IL}\mathsf{M}_0$–frame that satisfies all the invariants, possibly $\mathcal{I}_{\mathsf{M}_0}$ excluded, then we can assume, nevertheless, that $\mathcal{I}_{\mathsf{M}_0}$ holds as well.

Corollary 8.20. *Any quasi–$\mathbf{IL}\mathsf{M}_0$–frame that satisfies all of the above invariants, except possibly $\mathcal{I}_{\mathsf{M}_0}$, can be extended to an $\mathbf{IL}\mathsf{M}_0$–frame that satisfies all the invariants.*

Proof. Only $\mathcal{I}_{\mathcal{D}}$ and $\mathcal{I}_{\mathrm{d}}$ need some attention. All the other invariants are given in terms of relations that do not change during the construction of the $\mathbf{IL}\mathsf{M}_0$-closure (Lemma 8.13). $\dashv$

Lemma 8.21. *Let $F = \langle W, R, S, \nu\rangle$ be a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. Then $F \models x \in \mathcal{N}_w^A$ iff. one of the following cases applies.*

[17] We call them sub–invariants since they merely serve the purpose of showing that the main-invariants are, indeed, invariant.

1. $\nu(w,x) = A$

2. There exists $t \in \mathcal{N}_w^A$ *such that* tKx

3. There exists $t \in \mathcal{N}_w^A$ *such that* $t\mathcal{S}_w x$

Corollary 8.22. *Let F be a quasi–*$\mathbf{IL}\mathsf{M}_0$*–frame that satisfies $\mathcal{J}_{\nu_4}$. Let $w, x \in F$ and let A be a formula. Then $x \in \mathcal{N}_w^A$ implies $\nu(w,x) = A$ or there exists some $t \in \mathcal{N}_w^A$ such that tKx.*

Lemma 8.23. *Let F be a quasi–frame which satisfies $\mathcal{J}_{\mathcal{N}_2}$, $\mathcal{J}_{\nu_1}$, $\mathcal{J}_{\nu_3}$ and $\mathcal{J}_{\nu_4}$. Then $x\mathcal{S}_v y$, $y \in \mathcal{N}_w^A \Rightarrow x \in \mathcal{N}_w^A$.*

Proof. Suppose $x\mathcal{S}_v y$ and $y \in \mathcal{N}_w^A$. Then, by Corollary 8.22, $\nu(w,y) = A$ or, for some $t \in \mathcal{N}_w^A$, tKy. In the first case we obtain $w = v$ by $\mathcal{J}_{\nu_3}$ and $\mathcal{J}_{\nu_4}$. And thus by $\mathcal{J}_{\mathcal{N}_2}$, $x \in \mathcal{N}_w^A$. In the second case we have, by $\mathcal{J}_{\nu_4}$ and $\mathcal{J}_{\nu_1}$ that $t \subset v$. Which implies, by Lemma 8.18–1., tKx. ⊣

8.5 Solving problems

Let $F = \langle W, R, S, \nu \rangle$ be a quasi–$\mathbf{IL}\mathsf{M}_0$–frame that satisfies all the invariants. Let $(\mathbf{a}, \neg(A \rhd B))$ be a $\mathcal{D}$-problem in F. We fix some $\mathbf{b} \notin W$. Using Lemma 4.17 we find a MCS $\Delta_\mathbf{b}$, such that $\nu(\mathbf{a}) \prec_B \Delta_\mathbf{b}$ and $A, \Box\neg A \in \Delta_\mathbf{b}$. We put

$$\begin{aligned}\hat{F} &= \langle \hat{W}, \hat{R}, \hat{S}, \hat{\nu} \rangle \\ &= \langle W \cup \{\mathbf{b}\}, R \cup \{\langle \mathbf{a}, \mathbf{b} \rangle\}, S, \nu \cup \{\langle \mathbf{b}, \Delta_\mathbf{b} \rangle, \langle \langle \mathbf{a}, \mathbf{b} \rangle, B \rangle\} \rangle,\end{aligned}$$

and define $\hat{K} = K(\hat{F})$. The frames F and $\hat{F}$ satisfy the conditions of Lemma 8.15. Thus we have

$$\forall xy{\in}F \; xKy \Leftrightarrow x\hat{K}y. \tag{14}$$

Since $\hat{S}{=}S$, this implies that all simple enough properties expressed in $\hat{K}$ and $\hat{S}$ using only parameters from F are true if they are true with $\hat{K}$ replaced by K.

Claim. *$\hat{F}$ is a quasi–*$\mathbf{IL}\mathsf{M}_0$*–frame.*

Proof. A simple check of Properties (1.–5.) of Definition 5.1 (quasi–frames) and Properties (6.–10.) of Definition 8.10 (quasi–$\mathbf{IL}\mathsf{M}_0$–frames) and the remaining ones in Definition 5.1 (quasi–frames). Let us comment on two of them.

$x\hat{K}y \to \hat{\nu}(x) \prec \hat{\nu}(y)$ follows from Lemma 8.14 and (14).

Let us show $\hat{F} \models x \in \mathcal{N}_w^C \Rightarrow \hat{\nu}(w) \prec_C \hat{\nu}(x)$. We have $\forall xw{\in}F \; F \models x \in \mathcal{N}_w^C \Leftrightarrow \hat{F} \models x \in \mathcal{N}_w^C$. So we only have to consider the case $\hat{F} \models \mathbf{b} \in \mathcal{N}_w^C$. If $w = \mathbf{a}$ then we are done by choice of $\hat{\nu}(\mathbf{b})$. Otherwise, by Lemma 8.23, we have for some $x \in F$, $F \models x \in \mathcal{N}_w^C$ and $x\hat{K}\mathbf{b}$. By the first property we proved, we get $\hat{\nu}(x) \prec \hat{\nu}(\mathbf{b})$. So, since $\hat{\nu}(w) \prec_C \hat{\nu}(x)$ we have $\hat{\nu}(w) \prec_C \hat{\nu}(\mathbf{b})$. ⊣

Before we show that $\hat{F}$ satisfies all the invariants we prove some lemmata.

Lemma 8.24. *If for some $x \neq \mathbf{a}$, $x\hat{K}^1\mathbf{b}$. Then there exist unique u and w (independent of x) such that $wK^{\geq 2}u(\mathcal{S}_w)^{tr}\mathbf{a}$.*

Proof. If such w and u do not exists then $T = K \cup \{\mathbf{a}, \mathbf{b}\}$ satisfies the conditions of Lemma 8.8. In which case $xK^1\mathbf{b}$ gives $xT\mathbf{b}$ which implies $x = \mathbf{a}$. The uniqueness of w follows from $\mathcal{J}_{\nu_3}$ and $\mathcal{J}_{\nu_4}$. The uniqueness of u follows from $\mathcal{J}_{\mathrm{u}}$ and the uniqueness of w. ⊣

In what follows we will denote these w and u, if they exist, by $\mathbf{w}$ and $\mathbf{u}$.

Lemma 8.25. *For all x. If $x\hat{K}^1\mathbf{b}$ then $x \subset \mathbf{a}$.*

Proof. Let $K' = K \cup \{(x, \mathbf{b}) \mid x\hat{K}\mathbf{b} \wedge x \subset \mathbf{a}\}$. It is not hard to show that K' satisfies the conditions of T in Lemma 8.8. ⊣

Lemma 8.26. *Suppose the conditions of Lemma 8.24 are satisfied and let $\mathbf{u}$ be the u asserted to exist. Then for all $x \neq \mathbf{a}$, if $x\hat{K}^1\mathbf{b}$, then $xK^1\mathbf{u}$.*

Proof. By Lemma 8.25 we have $x \subset \mathbf{a}$. Let

$$x = x_0(\subset_1 \cup K)x_1(\subset_1 \cup K)\cdots(\subset_1 \cup K)x_n = \mathbf{a}.$$

First we show $x = x_0 \subset_1 x_1 \subset_1 \cdots \subset_1 x_n = \mathbf{a}$. Suppose, for a contradiction, that for some $i < n$, x_iKx_{i+1}. Then, by Lemma 8.18, $xKx_{i+1}K\mathbf{b}$. So, $xK^{\geq 2}\mathbf{b}$. A contradiction. The lemma now follows by showing, with induction on i and using $F \models \mathcal{J}_{K^1}$, that for all $i \geq 0$, $x_{n-(i+1)}K^1\mathbf{u}$.
⊣

Lemma. *$\hat{F}$ satisfies all the sub-invariants.*

Proof. We only comment on $\mathcal{J}_{K^1}$ and $\mathcal{J}_{\nu_1}$. Let $K = K(\hat{F})$.

$\mathcal{J}_{\nu_1}$ follows from Lemma 8.25, so let us treat $\mathcal{J}_{K^1}$. Suppose $w\hat{K}x\hat{K}^1y(\hat{\mathcal{S}}_w)^{\mathrm{tr}}y'\hat{K}^1z$. We can assume that at least one of w, x, y, y', z is not in F and the only candidate for this is z. So we have $z = \mathbf{b}$. We can assume that $x \neq y'$ (otherwise we are done at once), so the conditions of Lemma 8.24 are fulfilled and thus $\mathbf{w}$ and $\mathbf{u}$ as stated there exist.

Suppose now, for a contradiction, that for some t, $x\hat{K}t\hat{K}^1\mathbf{b}$. Then by Lemma 8.26, $t = \mathbf{a}$ or $t\hat{K}^1\mathbf{u}$. Suppose we are in the case $t = \mathbf{a}$. Since $\nu(\mathbf{w}, \mathbf{a})$ is defined and $x\hat{K}\mathbf{a}$ we obtain by $\mathcal{J}_{\nu_1}$, that $x \subset \mathbf{w}$. Since $\mathbf{w}\hat{K}^{\geq 2}\mathbf{u}$ we obtain by Lemma 8.18 that $x\hat{K}^{\geq 2}\mathbf{u}$. In the case $t\hat{K}^1\mathbf{u}$ we have $x\hat{K}^{\geq 2}\mathbf{u}$ trivially. So in any case we have

$$x\hat{K}^{\geq 2}\mathbf{u}.$$

However, by Lemma 8.26 and since $y'\hat{K}^1z$ we have $y'\hat{K}^1\mathbf{u}$ or $y' = \mathbf{a}$. In the first case, since $F \models \mathcal{J}_{K^1}$, we have $x\hat{K}^1\mathbf{u}$. In the second case we obtain, by the uniqueness of $\mathbf{u}$, that $y = \mathbf{u}$ and thus $x\hat{K}^1\mathbf{u}$. So in any case we have

$$x\hat{K}^1\mathbf{u}.$$

A contradiction.
⊣

Lemma. *Possibly with the exception of* $\mathcal{I}_{\mathsf{M}_0}$, $\hat{F}$ *satisfies all the main-invariants.*

Proof. Let $K = K(\hat{F})$. We only comment on $\mathcal{I}_\Box$ and $\mathcal{I}_\mathcal{N}$.

First we treat $\mathcal{I}_\Box$. So we have to show that for all y, $\{\hat{\nu}(x) \mid x\hat{K}^1 y\}$ is linearly ordered by $\subseteq_\Box$. We only need to consider the case $y = \mathbf{b}$. If $\{\mathbf{a}\} = \{x \mid x\hat{K}^1\mathbf{b}\}$ then the claim is obvious. So we can assume that the condition of Lemma 8.24 is fulfilled and we fix $\mathbf{u}$ as stated. The claim now follows by $F \models \mathcal{I}_\Box$ (with $y = \mathbf{u}$) and noting that, by Lemma 8.14, $x\hat{K}^1\mathbf{b} \Rightarrow x \subseteq_\Box \mathbf{a}$.

Now we look at $\mathcal{I}_\mathcal{N}$: $w\hat{K}x\hat{K}y \wedge \hat{F} \models y \in \mathcal{N}^A_w \to \hat{F} \models x \in \mathcal{N}^A_w$. Suppose $w\hat{K}x\hat{K}y$ and $\hat{F} \models y \in \mathcal{N}^A_w$. We only have to consider the case $y = \mathbf{b}$. Then, by Lemma 8.21, $\hat{\nu}(w, \mathbf{b}) = A$ or for some $t \in \mathcal{N}^A_w$ we have $t\hat{\mathcal{S}}_w\mathbf{b}$ or $t\hat{K}^1\mathbf{b}$. The first case is impossible by $\mathcal{J}_{\nu_2}$. The second is also clearly not so. Thus we have

$$t\hat{K}^1\mathbf{b}. \tag{15}$$

We suppose that the conditions of Lemma 8.24 are fulfilled (the other case is easy). If $t\hat{K}^1\mathbf{u}$ and $x\hat{K}^*\mathbf{u}$ then we are done simmilarly as the case above. So assume $t\hat{K}^1\mathbf{a}$ or $x\hat{K}^*\mathbf{a}$. Since wRt and wRx in any case we have $w\hat{K}\mathbf{a}$. Now by Lemma 8.23 and $\mathcal{J}_{\mathcal{N}_1}$ we have $\mathbf{u} \in \mathcal{N}^A_w \Leftrightarrow \mathbf{a} \in \mathcal{N}^A_w$. Also, by (15), $\mathbf{u} \in \mathcal{N}^A_w \vee \mathbf{a} \in \mathcal{N}^A_w$. So since $x\hat{K}\mathbf{u}$ or $x = \mathbf{a}$ or $x\hat{K}\mathbf{a}$ we obtain $x \in \mathcal{N}^A_w$ by $F \models \mathcal{I}_\mathcal{N}$.

$\dashv$

To finish this subsection we note that by Lemma 8.13 and Corollary 8.20 we can extend $\hat{F}$ to an adequate $\mathbf{IL}\mathsf{M}_0$–frame that satisfies all invariants.

8.6 Solving deficiencies

Let $F = \langle W, R, S, \nu\rangle$ be an $\mathbf{IL}\mathsf{M}_0$–frame satisfing all the invariants. Let $(\mathbf{a}, \mathbf{b}, C \rhd D)$ be a $\mathcal{D}$-deficiency in F.

Suppose $\mathbf{a}R^{\geq 2}\mathbf{b}$ (the case $\mathbf{a}R^1\mathbf{b}$ is easy). Let x be the $\subseteq_\Box$-maximum of $\{x \mid \mathbf{a}KxK^1\mathbf{b}\}$. This maximum exists by $\mathcal{I}_\Box$. Pick some A such that $\mathbf{b} \in \mathcal{N}^A_\mathbf{a}$. (If such an A exists, then by adequacy of F, it is unique. If no such A exists, take $A = \bot$.) By $\mathcal{I}_\mathcal{N}$ and adequacy we have $\nu(\mathbf{a}) \prec_A \nu(x)$. So we have $C \rhd D \in \nu(\mathbf{a}) \prec_A \nu(x) \ni \Diamond C$. We apply Lemma 8.16 to obtain, for some set Y, disjoint from W, a set $\{\Delta_y \mid y \in Y\}$ of MCS's with all the properties as stated in that lemma. We define

$$\begin{aligned}\hat{F} = \langle W \cup Y, &R \cup \{\langle \mathbf{a}, y\rangle \mid y \in Y\},\\ &S \cup \{\langle \mathbf{a}, \mathbf{b}, y\rangle \mid y \in Y\} \cup \{\langle \mathbf{a}, y, y'\rangle \mid y, y' \in Y, y \neq y'\},\\ &\nu \cup \{\langle y, \Delta_y\rangle, \langle\langle \mathbf{a}, y\rangle, A\rangle \mid y \in Y\}\rangle.\end{aligned}$$

Claim. $\hat{F}$ *is a quasi–*$\mathbf{IL}\mathsf{M}_0$*–frame.*

Proof. An easy check of Properties (1.–5.) of Definition 5.1 (quasi–frames) and Properties (6.–10.) of Definition 8.10 (quasi–$\mathbf{IL}\mathsf{M}_0$–frames). Let us comment on two cases.

First we see that $x\hat{K}y \to \hat{\nu}(x) \prec \hat{\nu}(y)$. We can assume $y \in Y$. By Lemma 8.14 we obtain some z with $\hat{\nu}(x) \subseteq_\Box \hat{\nu}(z)$ and $x(\hat{R} \cup \hat{S})^* z\hat{R}y$. This z can only be $\mathbf{a}$. By choice of $\hat{\nu}(y)$ we have $\hat{\nu}(\mathbf{a}) \prec \hat{\nu}(y)$. And thus $\hat{\nu}(x) \prec \hat{\nu}(y)$.

We now see that $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* y' \to \hat{\nu}(x) \subseteq_\Box \hat{\nu}(y')$. We can assume at least one of w, x, y, y' is in Y. The only candidates for this are y and y'. If both are in Y then $w = \mathbf{a}$ and an x as stated does not exists. So only $y' \in Y$ and thus in particular $y \neq y'$. Now there are two cases to consider.

The first case is that for some t, $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* t\hat{K}y'$. But, $\hat{\nu}(y')$ is $\subseteq_\Box$-larger than $\hat{\nu}(t)$ by $x\hat{K}y \to \hat{\nu}(x) \prec \hat{\nu}(y)$. Also we have $wKxKy(S_w \cup K)^* t$. So, $\hat{\nu}(x) = \nu(x) \subseteq_\Box \nu(t) = \hat{\nu}(t)$.

The second case is $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* \mathbf{b}\hat{S}_w y'$. In this case we have $w = \mathbf{a}$. y' is chosen to be $\subseteq_\Box$–larger than the $\subseteq_\Box$-maximum of $\{\nu(r) \mid \mathbf{a}KrK^1\mathbf{b}\}$. We have $wKxKy(S_w \cup K)^* \mathbf{b}$ So, by $F \models \mathcal{I}_S$, this $\subseteq_\Box$–maximum is $\subseteq_\Box$–larger than $\nu(x)$. ⊣

Lemma 8.27. *For any $x \in \hat{F}$ and $y \in Y$ we have $x\hat{K}^1 y \to x \subset \mathbf{a}$.*

Proof. We put $K' = K \cup \{(x, y) \mid y \in Y,\ x\hat{K}y,\ x \subset \mathbf{a}\}$. By showing that K' satisfies the conditions of T in Lemma 8.8. we obtain $x\hat{K}^1 y \to xK'y$. So if $x\hat{K}^1 y$ then $xK'y$. But if $y \in Y$ then xKy does not hold. Thus we have $x \subset \mathbf{a}$. ⊣

Lemma 8.28. *Suppose $y \in Y$ and $\mathbf{a}\hat{K}^1 z$. Then for all x, $x\hat{K}^1 y \to x\hat{K}^1 z$.*

Proof. Suppose $xK^1 y$. By Lemma 8.27 we have $x \subset \mathbf{a}$. There exist $x_0, x_1, x_2, \ldots, x_n$ such that $x = x_0(\subset_1 \cup K)x_1(\subset_1 \cup K)\cdots(\subset_1 \cup K)x_n = \mathbf{a}$. First we show that $x = x_0 \subset_1 x_1 \subset_1 \cdots \subset_1 \mathbf{a}$. Suppose, for a contradiction that for some $i < n$, we have $x_i K x_{i+1}$. Then $xKx_{i+1}Ky$ and thus $xK^{\geq 2}y$. A contradiction. The lemma now follows by showing, with induction on i, using $\mathcal{J}_{K^1}$, that for all $i \leq n$, $x_{n-i}K^1 z$. ⊣

Lemma. *$\hat{F}$ satisfies all the sub-invariants.*

Proof. The proofs are rather straightforward. We give two examples.

First we show $\mathcal{J}_{\mathrm{u}}$: $w\hat{K}^{\geq 2}x(\hat{\mathcal{S}}_w)^{tr} y \wedge w\hat{K}^{\geq 2}x'(\hat{\mathcal{S}}_w)^{tr} y \to x = x'$. Suppose $w\hat{K}^{\geq 2}x(\hat{\mathcal{S}}_w)^{tr} y$ and $w\hat{K}^{\geq 2}x'(\hat{\mathcal{S}}_w)^{tr} y$. We can assume that $y \in Y$. (Otherwise all of w, x, x', y are in F and we are done by $F \models \mathcal{J}_{\mathrm{u}}$.) We clearly have $w \in F$. If $x \in Y$ then $w = \mathbf{a}$ and thus $w\hat{K}^1 x$. So, $x \notin Y$. Next we show that both $x, x' \neq \mathbf{b}$.

Assume, for a contradiction, that at least one of them equals $\mathbf{b}$. W.l.o.g. we assume it is x. But then $wK^{\geq 2}\mathbf{b}$ and $wK^{\geq 2}x'(\mathcal{S}_w)^{tr}\mathbf{b}$. By $F \models \mathcal{J}_{\nu_4}$ we now obtain that $\nu(w, \mathbf{b})$ is defined. And thus by $F \models \mathcal{J}_{\nu_2}$, $wK^1\mathbf{b}$. A contradiction.

So, both $x, x' \neq \mathbf{b}$. But now $wK^{\geq 2}x(\mathcal{S}_w)^{tr}\mathbf{b}$ and $wK^{\geq 2}x'(\mathcal{S}_w)^{tr}\mathbf{b}$. So, by $F \models \mathcal{J}_{\mathrm{u}}$, we obtain $x = x'$.

Now let us see that $\mathcal{J}_{K^1}$ holds, that is $w\hat{K}x\hat{K}^1 y(\hat{\mathcal{S}}_w)^{tr} y'\hat{K}^1 z \to x\hat{K}^1 z$. Suppose $w\hat{K}x\hat{K}^1 y(\hat{\mathcal{S}}_w)^{tr} y'\hat{K}^1 z$. We can assume that $z \in Y$. (Otherwise all of w, x, y, y', z are in F and we are done by $F \models \mathcal{J}_{K^1}$.) Fix some $a_1 \in F$ for which $\mathbf{a}K^1 a_1$. By Lemma 8.28 we have $y'K^1 a_1$ and thus, since $F \models \mathcal{J}_{K^1}$, $xK^1 a_1$.

By definition of $\hat{K}$ we have $x\hat{K}z$. Now, if for some t, we have $x\hat{K}t\hat{K}^1z$, then similarly as above,tK^1a_1. So, this implies $xK^{\geq 2}a_1$. A contradiction, conclusion: xK^1z. $\dashv$

Lemma. *Except for $\mathcal{I}_{\mathsf{M}_0}$, $\hat{F}$ satisfies all main-invariants.*

Proof. We only comment on $\mathcal{I}_\Box$ and $\mathcal{I}_\mathcal{N}$.

First we show $\mathcal{I}_\Box$: For all y, $\{\hat{\nu}(x) \mid x\hat{K}^1y\}$ is linearly ordered by $\subseteq_\Box$. Let $y \in \hat{F}$ and consider the set $\{x \mid xK^1y\}$. Since $\hat{K} \restriction_F = K$ and for all $y \in Y$ there does not exists z with $y\hat{K}^1z$ we only have to consider the case $y \in Y$. Fix some a_1 such that $\mathbf{a}K^1a_1K^*\mathbf{b}$. By Lemma 8.27 for any such y we have

$$\{x \mid xK^1y\} \subseteq \{x \mid xK^1a_1\}.$$

And by $F \models \mathcal{I}_\Box$ with a_1 for y, we know that $\{\nu(x) \mid xK^1a_1\}$ is linearly ordered by $\subseteq_\Box$.

Now let us see $\mathcal{I}_\mathcal{N}$: $w\hat{K}x\hat{K}y \wedge \hat{F} \models y \in \mathcal{N}^A_w \rightarrow \hat{F} \models x \in \mathcal{N}^A_w$. Suppose $w\hat{K}x\hat{K}y$ $\hat{F} \models y \in \mathcal{N}^A_w$. We can assume $y \in Y$. By Lemma 8.27, $x \subset \mathbf{a}$. So, $wKxK\mathbf{b}$. By Lemma 8.23, $F \models \mathbf{b} \in \mathcal{N}^A_w$ and thus $\hat{F} \models x \in \mathcal{N}^A_w$. $\dashv$

To finish this section we noting that by Lemma 8.13 and Corollary 8.20 we can extend $\hat{F}$ to an adequate $\mathbf{IL}\mathsf{M}_0$–frame that satisfies all invariants.

8.7 Rounding up

It is clear that the union of a bounded chain of $\mathbf{IL}\mathsf{M}_0$–frames is itself an $\mathbf{IL}\mathsf{M}_0$–frame.

9 The logic ILW*

In this section we are going to prove the following theorem.

Theorem 9.1. $\mathbf{IL}\mathsf{W}^*$ *is a complete logic.*

The completeness proof of $\mathbf{IL}\mathsf{W}^*$ lifts almost completely along with the completeness proof for $\mathbf{IL}\mathsf{M}_0$. We only need some minor adaptations.

9.1 Preliminaries

The frame condition of W is well known.

Theorem 9.2. *For any* $\mathbf{IL}$*-frame F we have that $F \models \mathsf{W} \Leftrightarrow \forall w\ (S_w; R)$ is conversely well-founded.*

In [dJV99] a completeness proof for $\mathbf{IL}\mathsf{W}$ was given. We can define a new principle M_0^* that is equivalent to W^*, as follows.

$$\mathsf{M}_0^* : \quad A \rhd B \rightarrow \Diamond A \wedge \Box C \rhd B \wedge \Box C \wedge \Box \neg A$$

Lemma 9.3. $\mathbf{IL}\mathsf{M_0W} = \mathbf{IL}\mathsf{W}^* = \mathbf{IL}\mathsf{M}_0^*$

Proof. The proof we give consists of four natural parts.

First we see $\mathbf{IL}\mathsf{W}^* \vdash \mathsf{M}_0$. We reason in $\mathbf{IL}\mathsf{W}^*$ and assume $A \rhd B$. Thus, also $A \rhd (B \vee \Diamond A)$. Applying the W^* axiom to the latter yields $(B \vee \Diamond A) \wedge \Box C \rhd (B \vee \Diamond A) \wedge \Box C \wedge \Box\neg A$. From this we may conclude

$$\begin{array}{lll} \Diamond A \wedge \Box C & \rhd & (B \vee \Diamond A) \wedge \Box C \\ & \rhd & (B \vee \Diamond A) \wedge \Box C \wedge \Box\neg A \\ & \rhd & B \wedge \Box C \end{array}$$

Secondly, we see that $\mathbf{IL}\mathsf{W}^* \vdash \mathsf{W}$. Again, we reason in $\mathbf{IL}\mathsf{W}^*$. We assume $A \rhd B$ and take the C in the W^* axiom to be $\top$. Then we immediately see that $A \rhd B \rhd B \wedge \Box\top \rhd B \wedge \Box\top \wedge \Box\neg A \rhd B \wedge \Box\neg A$.

We now easily see that $\mathbf{IL}\mathsf{M_0W} \vdash \mathsf{M}_0^*$. For, reason in $\mathbf{IL}\mathsf{M_0W}$ as follows. By W^*, $A \rhd B \rhd B \wedge \Box\neg A$. Now an application of M_0 on $A \rhd B \wedge \Box\neg A$ yields $\Diamond A \wedge \Box C \rhd B \wedge \Box C \wedge \Box\neg A$.

Finally we see that $\mathbf{IL}\mathsf{M}_0^* \vdash \mathsf{W}^*$. So, we reason in $\mathbf{IL}\mathsf{M}_0^*$ and assume $A \rhd B$. Thus, we have also $\Diamond A \wedge \Box C \rhd B \wedge \Box C \wedge \Box\neg A$. We now conclude $B \wedge \Box C \rhd B \wedge \Box C \wedge \Box\neg A$ easily as follows. $B \wedge \Box C \rhd (B \wedge \Box C \wedge \Box\neg A) \vee (\Box C \wedge \Diamond A) \rhd B \wedge \Box C \wedge \Box\neg A$. ⊣

Corollary 9.4. *For any* **IL***-frame we have that* $F \models \mathsf{W}^*$ *iff. both (for each* w, $(S_w; R)$ *is conversely well-founded) and* $(\forall w, x, y, y', z\ (wRxRyS_wy'Rz \rightarrow xRz))$.

The frame condition of W^* tells us how to correctly define the notions of adequate $\mathbf{IL}\mathsf{W}^*$-frames and quasi-$\mathbf{IL}\mathsf{W}^*$-frames.

Definition 9.5 ($\subsetneq_{\Box}^{\mathcal{D}}$). Let $\mathcal{D}$ be a finite set of formulas. Let $\subsetneq_{\Box}^{\mathcal{D}}$ be a binary relation on MCS's defined as follows. $\Delta \subsetneq_{\Box}^{\mathcal{D}} \Delta'$ iff.

1. $\Delta \subseteq_{\Box} \Delta'$,
2. For some $\Box A \in \mathcal{D}$ we have $\Box A \in \Delta' - \Delta$.

Lemma 9.6. *Let* F *be a quasi-frame and* $\mathcal{D}$ *be a finite set of formulas. If* $wRxRyS_wy' \rightarrow \nu(x) \subsetneq_{\Box}^{\mathcal{D}} \nu(y')$ *then* $(R; S_w)$ *is conversely well-founded.*

Proof. By the finiteness of $\mathcal{D}$. ⊣

Lemma 9.7. *Let* F *be a quasi-*$\mathbf{IL}\mathsf{M}_0$*-frame. If* $wRxRyS_wy' \rightarrow \nu(x) \subsetneq_{\Box}^{\mathcal{D}} \nu(y')$ *then* $wRxRy(S_w \cup R)^*y' \rightarrow \nu(x) \subsetneq_{\Box}^{\mathcal{D}} \nu(y')$

Proof. Suppose $wRxRy(S_w \cup R)^*y'$. $\nu(x) \subsetneq_{\Box}^{\mathcal{D}} \nu(y')$ follows with induction on the minimal number of R-steps in the path from y to y'. ⊣

Definition 9.8 (Adequate $\mathbf{IL}\mathsf{W}^*$-frame). Let $\mathcal{D}$ be a set of formulas. We say that an adequate $\mathbf{IL}\mathsf{M}_0$-frame is an adequate $\mathbf{IL}\mathsf{W}^*$-frame (w.r.t. $\mathcal{D}$) iff. the following additional property holds.

8. $wRxRy(\mathcal{S}_w)^{\mathrm{tr}}y' \rightarrow x \subsetneq_{\Box}^{\mathcal{D}} y'$

Definition 9.9 (Quasi-ILW*-frame). Let $\mathcal{D}$ be a set of formulas. We say that a quasi-**IL**M_0-frame is a quasi-**IL**W^*-frame (w.r.t. $\mathcal{D}$) iff. the following additional property holds.

13. $wKxKy(\mathcal{S}_w)^{\mathrm{tr}}y' \rightarrow x \subsetneq^{\mathcal{D}}_{\Box} y'$

In what follows we might simply talk of adequate **IL**W^*-frames and quasi-**IL**W^* In these cases $\mathcal{D}$ is clear from context.

Lemma 9.10. *Any quasi-***IL**W^**-frame can be extended to an adequate* **IL**W^**-frame. (Both w.r.t. the same set of formulas $\mathcal{D}$.)*

Proof. Let F be a quasi-**IL**W^*-frame. Then in particular F is a quasi-**IL**M_0-frame. So consider the proof of Lemma 8.13. There we constructed a sequence of quasi-**IL**M_0-frames $F = F_0 \subseteq F_1 \subseteq \bigcup_{i<\omega} F_i = \hat{F}$. What we have to do, is to show that if $F_0(= F)$ is a quasi-**IL**W^*-frame, then each F_i is as well. Additionally we have to show that $\hat{F}$ is an adequate **IL**W^*-frame.

But this is rather trivial. As noted in the proof of Lemma 8.13, The relation K and the relations $(\mathcal{S}_w)^{\mathrm{tr}}$ are constant throughout the whole process. So clearly each F_i is a quasi-**IL**W^*-frame.

Also the extra property of quasi-**IL**W^*-frames is preserved under unions of bounded chains. So, $\hat{F}$ is an adequate **IL**W^*-frame. $\dashv$

Lemma 9.11. *Let Γ and Δ be* MCS*'s with $\Gamma \prec_C \Delta$,*

$$P \rhd Q, S_1 \rhd T_1, \ldots, S_n \rhd T_n \in \Gamma \quad \textit{and} \quad \Diamond P \in \Delta.$$

There exist $k \leq n$. MCS*'s $\Delta_0, \Delta_1, \ldots, \Delta_k$ such that*

- *Each Δ_i lies C-critical above Γ,*
- *Each Δ_i lies $\subseteq_{\Box}$ above Δ,*
- *$Q \in \Delta_0$,*
- *For each $i \geq 0$, $\Box\neg P \in \Delta_i$,*
- *For all $1 \leq j \leq n$, $S_j \in \Delta_h \Rightarrow$ for some $i \leq k$, $T_j \in \Delta_i$.*

Proof. The proof is a straightforward adaptation of the proof of Lemma 8.16. In that proof, a trick was to postpone an application of M_0 as long as possible. We do the same here but let an application of M_0 on $P \rhd \Diamond P \vee \psi$ be preceded by an application of W to obtain $P \rhd \psi$. $\dashv$

9.2 Completeness

Again, we specify the four ingredients from Remark 4.19. The **Frame condition** is contained in Corollary 9.4.

The **Invariants** are all those of **IL**M_0 and additionally

$\mathcal{I}_{w^*}$ $wKxKy(\mathcal{S}_w)^{\mathrm{tr}}y' \rightarrow x \subsetneq^{\mathcal{D}}_{\Box} y'$

Here, $\mathcal{D}$ is some finite set of formulas closed under subformulas and single negation.

Problems. We have to show that we can solve problems in an adequate **IL**W^*-frame in such a way that we end up with a quasi-**IL**W^*-frame. If we have such a frame then in particular it is an **IL**$\mathsf{M_0}$-frame. So, as we have seen we can extend this frame to a quasi-**IL**$\mathsf{M_0}$-frame. It is easy to see that whenever we started with an adequate **IL**W^*-frame we end up with a quasi **IL**W^*-frame. (This is basically Lemma 9.10.)

Deficiencies. We have to show that we can solve any deficiency in an adequate **IL**W^*-frame such that we end up with an quasi-**IL**W^*-frame. It is easily seen that the process as described in the case of **IL**$\mathsf{M_0}$ works if we use Lemma 9.11 instead of Lemma 8.16.

Rounding up. We have to show that the union of a bounded chain of quasi-**IL**W^*-frames that satisfy all the invariants is an **IL**W^*-frame. The only novelty is that we have to show that in this union for each w we have that $(R; S_w)$ is conversely well-founded. But this is ensured by $\mathcal{I}_{w^*}$ and Lemma 9.6.

10 Incompleteness

10.1 Incompleteness of IL$\mathsf{P_0W^*}$

We shall now see the modal incompleteness of the logic **IL**$\mathsf{W^*P_0}$. We do this by showing that the principle R follows semantically from **IL**$\mathsf{W^*P_0}$ but is not provable in **IL**$\mathsf{W^*P_0}$.

Let us first calculate the frame condition of R. It turns out to be the same frame condition as for $\mathsf{P_0}$ (see [Joo98]).

Lemma 10.1. $F \models \mathsf{R} \Leftrightarrow [xRyRzS_xuRv \to zS_yv]$

Proof. "$\Leftarrow$" Suppose that at some world $x \Vdash A \rhd B$. We are to show $x \Vdash \neg(A \rhd \neg C) \rhd B \wedge \Box C$. Thus, if $xRy \Vdash \neg(A \rhd \neg C)$ we need to go via an S_x to a u with $u \vdash B \wedge \Box C$.

As $y \Vdash \neg(A \rhd \neg C)$, we can find z with $yRz \Vdash A$. Now, by $x \Vdash A \rhd B$, we can find u with $yS_xu \Vdash B$. We shall now see that $u \Vdash B \wedge \Box C$. For, if uRv, then by our assumption, zS_yv, and by $y \Vdash \neg(A \rhd \neg C)$, we must have $v \vdash C$. Thus, $u \Vdash B \wedge \Box C$ and clearly yS_xu.

"$\Rightarrow$" We suppose that R holds. Now we consider arbitrary a, b, c, d and e with $aRbRcS_adRe$. For propositional variables p, q and r we define a valuation $\Vdash$ as follows.

$$\begin{array}{lcl} x \Vdash p & :\Leftrightarrow & x = c \\ x \Vdash q & :\Leftrightarrow & x = d \\ x \Vdash r & :\Leftrightarrow & cS_bx \end{array}$$

Clearly, $a \Vdash p \rhd q$ and $b \Vdash \neg(p \rhd \neg r)$. By R we conclude $a \Vdash \neg(p \rhd \neg r) \rhd q \wedge \Box r$. Thus, $d \Vdash q \wedge \Box r$ which implies cS_be. ⊣

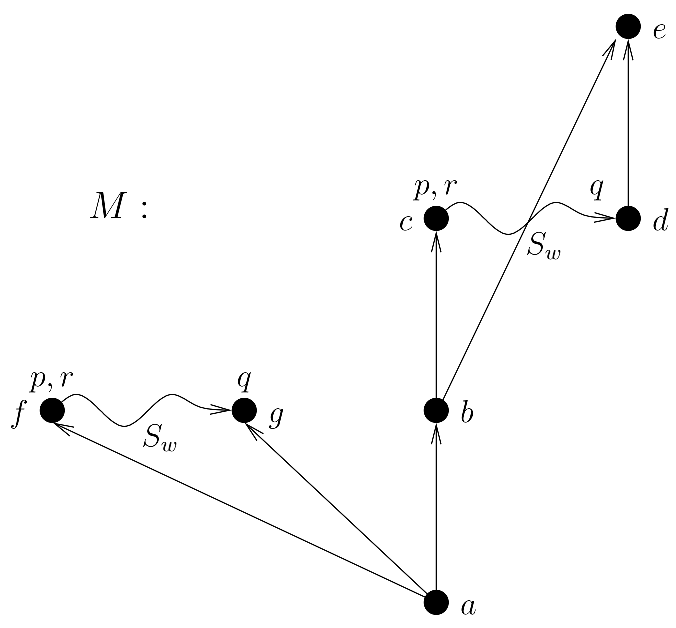

Figure 4: $\mathbf{IL}\mathsf{W}^*\mathsf{P}_0$ is incomplete

Theorem 10.2. $\mathbf{IL}\mathsf{W}^*\mathsf{P}_0 \nvdash \mathsf{R}$

Proof. We consider the model M from Figure 4 and shall see that $M \models \mathbf{IL}\mathsf{W}^*\mathsf{P}_0$ but $M, a \nVdash \mathsf{R}$. By Lemma 2.18 we conclude that $\mathbf{IL}\mathsf{W}^*\mathsf{P}_0 \nvdash \mathsf{R}$.

As M satisfies the frame condition for W^*, it is clear that $M \models \mathsf{W}^*$. We shall now see that $M \models A \rhd \Diamond B \to \Box(A \rhd B)$ for any formulas A and B.

A formula $\Box(A \rhd B)$ can only be false at some world with at least two successors. Thus, in M, we only need to consider the point a. So, supppose $A \rhd \Diamond B$. For which x with aRx can we have $x \Vdash A$?

As we have to be able to go via an S_x-transition to a world where $\Diamond B$ holds, the only candidates for x are b, c and d. But clearly, c and f make true the same modal formulas. From f it is impossible to go to a world where $\Diamond B$ holds.

Thus, if $a \Vdash A \rhd \Diamond B$, the A can only hold at b or at d. But this automatically implies that $a \Vdash \Box(A \rhd B)$ and $M \models \mathsf{P}_0$.

It is not hard to see that $a \nVdash \mathsf{R}$. Clearly, $a \Vdash p \rhd q$ and $b \Vdash \neg(p \rhd \neg r)$. However, $d \nVdash q \wedge \Box r$ and thus $a \nVdash \neg(p \rhd \neg r) \rhd q \wedge \Box r$. $\dashv$

The following lemma tells us that $\mathbf{IL}\mathsf{R}$ is a proper extension of $\mathbf{IL}\mathsf{M}_0\mathsf{P}_0$.

Lemma 10.3. $\mathbf{IL}\mathsf{R} \vdash \mathsf{M}_0, \mathsf{P}_0$

Proof. As $\mathbf{IL} \vdash \Diamond A \wedge \Box C \to \neg(A \rhd \neg C)$ we get that $A \rhd B \to \Diamond A \wedge \Box C \rhd \neg(A \rhd \neg C)$ and M_0 follows from R.

The principle P_0 follows directly from R by taking $C = \neg B$. $\dashv$

We can consider the principle R^* that can be seen, in a sense, as the union of W and R.

$$\mathsf{R}^* : \quad A \rhd B \to \neg(A \rhd \neg C) \rhd B \wedge \Box C \wedge \Box\neg A$$

Lemma 10.4. $\mathbf{IL}\mathsf{RW} = \mathbf{IL}\mathsf{R}^*$

Proof. $\supseteq$: $A \rhd B \to A \rhd B \wedge \Box\neg A \to \neg(A \rhd \neg C) \rhd B \wedge \Box C \wedge \Box\neg A$.

$\subseteq$: $A \rhd B \to \neg(A \rhd \neg C) \rhd B \wedge \Box C \wedge \Box\neg A \rhd B \wedge \Box C$; and if $A \rhd B$, then $A \rhd B \rhd ((B \wedge \Box\neg A) \vee \Diamond A) \rhd B \wedge \Box\neg A$, as $A \rhd B \to \neg(A \rhd \bot) \rhd B \wedge \Box\top \wedge \Box\neg A$. $\dashv$

10.2 Generalized semantics

In [Š91], Švejdar showed the independence of some extensions of **IL**. Some of these logics, however, had the same class of characteristic Veltman frames. Naturally, frames alone are not sufficient to distinguish between such logics so Švejdar used models combined with some bisimulation arguments instead. A generalized Veltman semantics, intended to uniformize this method, was proposed by de Jongh. This generalized semantics was previously investigated by Vuković [Vuk96], Joosten [Joo98] and Verbrugge and was successfully used to show independence of certain extensions of **IL**.

We will set both the generalized Veltman semantics and the model/bisimulation method to work in order to distinguish some extensions of **IL**, which are indistinguishable using Veltman frames alone. We use a slight variation of the semantics used in [Vuk96]. Any result in this section can be obtained with the old semantics, we think that nevertheless this might be a useful variation.

Definition 10.5 ($\mathbf{IL}_{\mathbf{set}}$-frame). A structure $\langle W, R, S\rangle$ is an $\mathbf{IL}_{\mathrm{set}}$-frame iff.

1. W is an non-empty set.
2. R is a transitive and conversely well-founded binary relation on W.
3. $S \subseteq W \times W \times (\mathcal{P}(W) - \{\emptyset\})$, such that (where we write yS_xY for $(x, y, Y) \in S$)
   (a) if xS_wY then wRx and for all $y \in Y$, wRy,
   (b) S is quasi-reflexive: wRx implies $xS_w\{x\}$,
   (c) S is quasi-transitive: If xS_wY then for all $y \in Y$ we have that if $y \notin Z$ and yS_wZ then xS_wZ,
   (d) $wRxRy$ implies $xS_w\{y\}$.

Definition 10.6 ($\mathbf{IL}_{\mathbf{set}}$-model). An $\mathbf{IL}_{\mathrm{set}}$-model is a structure $\langle W, R, S, \Vdash\rangle$ such that $\langle W, R, S\rangle$ is an $\mathbf{IL}_{\mathrm{set}}$-frame and $\Vdash$ is a binary relation between elements of W and modal formulas such that the following cases apply.

1. $\Vdash$ commutes with boolean connectives. For instance, $w \Vdash A \wedge B$ iff. $w \Vdash A$ and $w \Vdash B$.

2. $w \Vdash \Box A$ iff. for all x such that wRx we have that $x \Vdash A$.

3. $w \Vdash A \rhd B$ iff. for all x such that wRx and $x \Vdash A$ there exists some Y, such that xS_wY and for all $y \in Y$, $y \Vdash B$.

For $\mathbf{IL}_{\mathrm{set}}$-models $F = \langle W, R, S, \Vdash\rangle$ and $Y \subseteq W$ we will write $Y \Vdash A$ for $\forall y \in Y,\ y \Vdash A$.

As usual, we say that a formula A is valid on an $\mathbf{IL}_{\mathrm{set}}$-frame $F = \langle W, R, S\rangle$ if for any model $\overline{F} = \langle W, R, S, \Vdash\rangle$, based on F, and any $w \in W$, we have $\overline{F}, w \Vdash A$.

Lemma 10.7 (Soundness of IL). *If* $\mathbf{IL} \vdash A$ *then for any* $\mathbf{IL}_{set}$*-frame* F*,* $F \models A$*.*

Proof. Validity is preserved under modes pones and generalization and trivially any propositional tautology is valid on each $\mathbf{IL}_{\mathrm{set}}$-frame. So it is enough to show that all axioms of $\mathbf{IL}$ are valid on each $\mathbf{IL}_{\mathrm{set}}$-frame. We only treat $\mathsf{J2}$: $(A \rhd B) \wedge (B \rhd C) \to A \rhd C$.

Suppose $w \Vdash A \rhd B$ and $w \Vdash B \rhd C$. Pick some x with wRx and suppose $x \Vdash A$. There exists some Y with xS_wY and $Y \Vdash B$. W.l.o.g. we can assume that for some $y \in Y$, $y \nVdash C$. Fix such a y. Since $y \Vdash B$ and wRy there exists some Z such that yS_wZ and $Z \Vdash C$. In particular, $y \notin Z$. And thus we have xS_wZ. ⊣

Theorem 10.8 (Completeness of IL). *If* A *is valid on each* $\mathbf{IL}_{set}$*-frame, then* $\mathbf{IL} \vdash A$*.*

Proof. Suppose $\mathbf{IL} \nvdash A$. Then there exists an $\mathbf{IL}$-model $M = \langle W, R, S\rangle$, and some $m \in M$ such that $M, m \Vdash \neg A$. Let $M' = \langle W, R, S', \Vdash'\rangle$, where $\Vdash' = \Vdash$ on propositional variables and is extended as usual, and

$$S' = \{(w, x, Y) \mid \forall y \in Y\ xS_wy\}.$$

It is easy to see that M' is an $\mathbf{IL}_{\mathrm{set}}$-model. As an example let us see that S is quasi-transitive. Suppose xS'_wX, $y \in X$ and yS'_wY. (We can assume $y \notin Y$, but we won't use this.) Pick $y' \in Y$. Then xS_wy and yS_wy'. Thus xS_wy'. Since $y' \in Y$ was arbitrary we conclude xS'_wY.

A straightforward induction on B shows that for all B we have $w \Vdash' B \Leftrightarrow w \Vdash B$. Thus we have $m \Vdash' \neg A$ and in particular A is not valid on the underling frame of M'. ⊣

Definition 10.9 ($\mathbf{IL_{set}}\mathsf{M_0}$-frame). An $\mathbf{IL}_{\mathrm{set}}$-frame is an $\mathbf{IL}_{\mathrm{set}}\mathsf{M_0}$-frame iff. for all w, x, y, Y such that $wRxRyS_wY$ there exists some $Y' \subseteq Y$ such that

1. xS_wY' and

2. for all $y' \in Y'$ we have that for all z, $y'Rz \to xRz$.

Lemma 10.10. *For any* $\mathbf{IL}_{set}$*-frame* $F = \langle W, R, S\rangle$ *we have* $F \models \mathsf{M_0}$ *iff.* F *is an* $\mathbf{IL}_{set}\mathsf{M_0}$*-frame.*

Proof. ($\Leftarrow$) Suppose F is an $\mathbf{IL}_{\mathrm{set}}\mathsf{M}_0$-frame. Let $\overline{F} = \langle W, R, S, \Vdash\rangle$ be a model based on this frame. Pick $w \in W$ and suppose $w \Vdash A \rhd B$. Pick $x \in W$ with wRx and $x \Vdash \Diamond A \wedge \Box C$. Now there exists some y with xRy and $y \Vdash A$. Thus, for some Y, yS_wY and $Y \Vdash B$. Since F is an $\mathbf{IL}_{\mathrm{set}}\mathsf{M}_0$-frame, there exists some $Y' \subseteq Y$ such that xS_wY' and for all $y' \in Y'$ we have that for all z, $y'Rz \to xRz$. So, in particular, $Y' \Vdash \Box C$.

($\Rightarrow$) Suppose $F \models \mathsf{M}_0$. Choose w, x, y, Y such that $wRxRyS_wY$. Let p, q, s be distinct proposition variables. Define an $\mathbf{IL}_{\mathrm{set}}$-model $\overline{F} = \langle W, R, S, \Vdash\rangle$ as follows.

$$v \Vdash p \Leftrightarrow v = y$$
$$v \Vdash q \Leftrightarrow v \in Y$$
$$v \Vdash s \Leftrightarrow xRv$$

Now, $w \Vdash p \rhd q$ and thus $w \Vdash \Diamond p \wedge \Box s \rhd q \wedge \Box s$. Also, $x \Vdash \Diamond p \wedge \Box s$. So, there exists some Y' such that xS_wY' and $Y' \Vdash q \wedge \Box s$. But the only candidates for such an Y' are the subsets of Y. Also, since $Y' \Vdash \Box s$, by definition of $\Vdash$ we have $y' \in Y'$ and $y'Rz$ implies xRz. ⊣

Definition 10.11 ($\mathbf{IL}_{\mathbf{set}}\mathsf{P}_0$-frame). An $\mathbf{IL}_{\mathrm{set}}$-frame is an $\mathbf{IL}_{\mathrm{set}}\mathsf{P}_0$-frame iff. for all w, x, y, Y, Z such that

1. $wRxRyS_wY$ and
2. for all $y \in Y$ there exists some $z \in Z$ with yRz,

we have that there exists some $Z' \subseteq Z$ with yS_xZ'.

Lemma 10.12. *For any $\mathbf{IL}_{set}$-frame $F = \langle W, R, S\rangle$ we have $F \models \mathsf{P}_0$ iff. F is an $\mathbf{IL}_{set}\mathsf{P}_0$-frame.*

Proof. ($\Leftarrow$) Suppose F is an $\mathbf{IL}_{\mathrm{set}}\mathsf{P}_0$-frame. And let $\overline{F} = \langle W, R, S, \Vdash\rangle$ be an $\mathbf{IL}_{\mathrm{set}}$-model based on this frame. Let $w \in W$ and suppose $w \Vdash A \rhd \Diamond B$. Pick x, y in W with $wRxRy$ and $y \Vdash A$. There exists some Y with yS_wY and $Y \Vdash \Diamond B$. Put $Z = \{z \mid z \Vdash B\}$. Now for all $y \in Y$ there exists some $z \in Z$ such that yRz. So, there exists some $Z' \subseteq Z$ with yS_xZ'.

($\Rightarrow$) Suppose $F \models \mathsf{P}_0$. Choose $w, x, y \in W$ and $Y, Z \subseteq W$ such that $wRxRyS_wY$ and for all $y \in Y$ there exists some $z \in Z$ with yRz. Let p, q be distinct propositional variables. Define the $\mathbf{IL}_{\mathrm{set}}$-model $\overline{F} = \langle W, R, S, \Vdash\rangle$ as follows.

$$v \Vdash p \Leftrightarrow v = y$$
$$v \Vdash q \Leftrightarrow v \in Z$$

Now, $Y \Vdash \Diamond q$. So, $w \Vdash p \rhd \Diamond q$ and thus, since $w \Vdash \mathsf{P}_0$, $w \Vdash \Box(p \rhd q)$. So for some Z' we have yS_xZ and $Z' \Vdash q$. But the only candidates for such Z' are the subsets of Z. ⊣

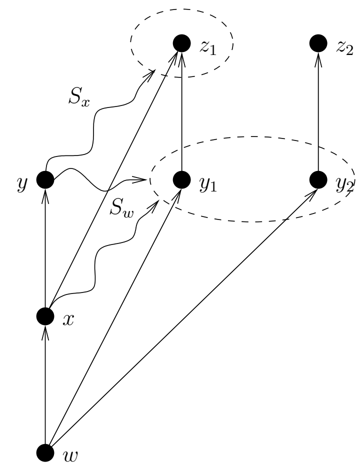

Figure 5: An $\mathbf{IL}_{\mathrm{set}}\mathsf{P}_0$-frame which is not an $\mathbf{IL}_{\mathrm{set}}\mathsf{M}_0$-frame.

Lemma 10.13. *There exists an* $\mathbf{IL}_{set}\mathsf{P}_0$*-frame which is not an* $\mathbf{IL}_{set}\mathsf{M}_0$*-frame.*

Proof. Consider Figure 5. It represents an $\mathbf{IL}_{\mathrm{set}}$-frame. For clarity we have omitted the following arrows. Those needed for the transitivity of R. Those needed for the quasi-reflexivity of S. Those needed for the inclusion of S in R. Additionally, quasi-transitivity dictates that we need $xS_w\{z_2\}$, $yS_w\{z_1\}$ and $yS_w\{z_2\}$. All the other ones are drawn.

Let us first see that we actually have an $\mathbf{IL}_{\mathrm{set}}\mathsf{P}_0$-frame. So suppose $vRaRbS_vB$. And let Z be such that for all $b' \in B$ there exists some $z \in Z$ such that $b'Rz$. It is not hard to see that only for $v = w$, $a = x$, $b = y$ and $B = \{y_1, y_2\}$ such a Z exists. And that moreover this Z must equal $\{z_1, z_2\}$. According to the P_0-condition we must find a $Z' \subseteq Z$ such that yS_xZ. And $\{z_1\}$ is such a Z'.

Now let us see that we do not have an $\mathbf{IL}_{\mathrm{set}}\mathsf{M}_0$-frame. Put $Y = \{y_1, y_2\}$. We have $wRxRyS_wY$. So, if we do have an $\mathbf{IL}_{\mathrm{set}}\mathsf{M}_0$-frame then for some $Y' \subseteq Y$ we have xS_wY' and for all $y' \in Y'$ we have that for all z, $y'Rz$ implies xRz. But the only $Y' \subseteq Y$ for which xS_wY' is Y itself. We have $y_2 \in Y$, y_2Rz_2 but not xRz_2. ⊣

Theorem 10.14. $\mathbf{IL}\mathsf{P}_0 \nvdash \mathsf{M}_0$.

Proof. If $\mathbf{IL}\mathsf{P}_0 \vdash \mathsf{M}_0$ then M_0 is valid on any $\mathbf{IL}_{\mathrm{set}}\mathsf{P}_0$-frame. But then any $\mathbf{IL}_{\mathrm{set}}\mathsf{P}_0$-frame is an $\mathbf{IL}_{\mathrm{set}}\mathsf{M}_0$-frame. Which, by Lemma 10.13 is not so. ⊣

Definition 10.15. Let $F = \langle W, R, S\rangle$ be an $\mathbf{IL}_{\mathrm{set}}$-frame. For any wRx we say that $\Gamma \subseteq W$ is a choice set for (w, x) iff. for all X such that xS_wX, $X \cap \Gamma \neq \emptyset$.

Definition 10.16. Let $F = \langle W, R, S\rangle$ be an $\mathbf{IL}_{\mathrm{set}}$-frame. We say that F is an $\mathbf{IL}_{\mathrm{set}}\mathsf{R}$-frame iff. $wRxRyS_wY$ implies that for all choice sets Γ for (x, y) there exists some $Y' = Y'(\Gamma) \subseteq Y$ such that xS_wY' and for all $y' \in Y'$ we have that for all z, $y'Rz$ implies $z \in \Gamma$.

Lemma 10.17. *An* $\mathbf{IL}_{set}$*-frame* $F = \langle W, R, S\rangle$ *is an* $\mathbf{IL}_{set}\mathsf{R}$*-frame iff.* $F \models \mathsf{R}$.

Proof. ($\Rightarrow$) Suppose F is an $\mathbf{IL}_{\mathrm{set}}\mathsf{R}$-frame. Let $\overline{F} = \langle W, R, S, \Vdash\rangle$ be a model based on F. Choose $w, x \in W$ and suppose wRx, $w \Vdash A \rhd B$ and $x \Vdash \neg(A \rhd C)$. We have to find some Y' with xS_wY' and $Y' \Vdash B \wedge \Box\neg C$. There exists some $y \in W$ such that xRy, $y \Vdash A$ and for all U such that yS_xU there exists some $u \in U$ with $u \Vdash \neg C$. Let Γ be a choice set for (x, y) such that $\Gamma \Vdash \neg C$ and $\Gamma \subseteq \bigcup_{yS_xU} U$. Since $w \Vdash A \rhd B$ we can find some Y such that yS_wY and $Y \Vdash B$. By the R frame condition we can find some $Y' \subseteq Y$ such that xS_wY' and for all $y' \in Y'$ we have that for all z, $y'Rz$ implies $z \in \Gamma$. So since $\Gamma \Vdash \neg C$ we conclude $Y' \Vdash B \wedge \Box\neg C$.

($\Leftarrow$). Suppose $F \models \mathsf{R}$. Let $w, x, y, Y \in W$ and suppose $wRxRyS_wY$. Let Γ be a choice set for (x, y). Let p, q, s, t be distinct propositional variables. Define the $\mathbf{IL}_{\mathrm{set}}$-model $\overline{F} = \langle W, R, S, \Vdash\rangle$ as follows.

$$v \Vdash p \Leftrightarrow v = y$$
$$v \Vdash q \Leftrightarrow v \in Y$$
$$v \Vdash s \Leftrightarrow v \notin \Gamma$$

Now, $w \Vdash p \rhd q$. So, $w \Vdash \neg(p \rhd s) \rhd q \wedge \Box\neg s$. Also, $\Gamma \Vdash \neg s$. So, $x \Vdash \neg(p \rhd s)$ and therefore there exists some Y' such that xS_wY' and $Y' \Vdash q \wedge \Box s$. Since $Y' \Vdash q$ we must have $Y' \subseteq Y$. Now let $y' \in Y'$ and pick some z for which $y'Rz$. Then $z \Vdash \neg s$ and thus by definition of $\Vdash$, $z \in \Gamma$. $\dashv$

References

[AdJH98] C. Areces, D. de Jongh, and E. Hoogland. The interpolation theorem for IL and ILP. In *Proceedings of AiML98. Advances in Modal Logic*, Uppsala. Sweden, October 1998. Uppsala University.

[Bek93] L.D. Beklemishev. On the complexity of arithmetic applications of modal formulae. *Archive for Mathematical Logic*, 32:229–238, 1993.

[Ber90] A. Berarducci. The interpretability logic of Peano arithmetic. *Journal of Symbolic Logic*, 55:1059–1089, 1990.

[Boo93] G. Boolos. *The Logic of Provability.* Cambridge University Press, Cambridge, 1993.

[BV01] M. Blackburn, P. Rijke and Y. Venema. *Modal Logic.* Camebridge University Press, 2001.

[dJJ98] D. de Jongh and G. Japaridze. The Logic of Provability. In S.R. Buss, editor, *Handbook of Proof Theory.* Studies in Logic and the Foundations of Mathematics, Vol.137., pages 475–546. Elsevier, Amsterdam, 1998.

[dJP96] D. de Jongh and D. Pianigiani. Solution of a problem of David Guaspari. *Studia Logica*, 1996.

[dJV90] D.H.J. de Jongh and F. Veltman. Provability logics for relative interpretability. In *[Pet90]*, pages 31–42, 1990.

[dJV91] D.H.J. de Jongh and A. Visser. Explicit fixed points in interpretability logic. *Studia Logica*, 50:39–50, 1991.

[dJV99] D.H.J. de Jongh and F. Veltman. Modal completeness of ILW. In J. Gerbrandy, M. Marx, M. Rijke, and Y. Venema, editors, *Essays dedicated to Johan van Benthem on the occasion of his 50th birthday*. Amsterdam University Press, Amsterdam, 1999.

[Gor03] E. Goris. Extending ILM with an operator for Σ_1–ness. Illc prepublication series, University of Amsterdam, 2003.

[Gua83] D. Guaspari. Sentences implying their own provability. *Journal of Symbolic Logic*, 48:777–789, 1983.

[HH02] R. Hirsch and I. Hodkinson. *Relation Algebras by Games*, volume 147 of *Studies in Logic*. Elsevier, North-Holland, 2002.

[HP93] P. Hájek and P. Pudlák. *Metamathematics of First Order Arithmetic*. Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[HV01] S. Hodkinson, I. Mikulás and Y. Venema. Axiomatizing complex algebras by games. *Algebra Universalis*, 46:455–478, 2001.

[Ign93] K.N. Ignatiev. The provability logic of Σ_1-interpolability. *Annals of Pure and Applied Logic*, 64:1–25, 1993.

[Jap94] G.K. Japaridze. The logic of the arithmetical hiearchy. *Anals of Pure and Applied Logic*, 66:89–112, 1994.

[Joo98] J.J. Joosten. Towards the interpretability logic of all reasonable arithmetical theories. Master's thesis, University of Amsterdam, 1998.

[JV00] J.J. Joosten and A. Visser. The interpretability logic of *all* reasonable arithmetical theories. *Erkenntnis*, 53(1–2):3–26, 2000.

[Ken73] C.F Kent. The relation of A to $Prov\ulcorner !A\urcorner$ in the Lindenbaum sentence algebra. *Journal of Symbolic Logic*, 38:359–367, 1973.

[Pet90] P.P. Petkov, editor. *Mathematical logic, Proceedings of the Heyting 1988 summer school in Varna, Bulgaria*. Plenum Press, Boston, 1990.

[Sha88] V. Shavrukov. The logic of relative interpretability over Peano arithmetic (in Russian). Technical Report Report No.5, Steklov Mathematical Institute, Moscow, 1988.

[Sol76] R.M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 28:33–71, 1976.

[TMR53] A. Tarski, A. Mostowski, and R. Robinson. *Undecidable theories.* North–Holland, Amsterdam, 1953.

[Š91] V. Švejdar. Some independence results in interpretability logic. *Studia Logica*, 50:29–38, 1991.

[Vis90] A. Visser. Interpretability logic. In *[Pet90]*, pages 175–209, 1990.

[Vis95] A. Visser. A course on bimodal provability logic. *Annals of Pure and Applied Logic*, pages 109–142, 1995.

[Vis97] A. Visser. An overview of interpretability logic. In M. Kracht, M. de Rijke, and H. Wansing, editors, *Advances in modal logic '96*, pages 307–359. CSLI Publications, Stanford, CA, 1997.

[Vis02] A. Visser. Faith & Falsity: a study of faithful interpretations and false Σ^0_1-sentences. Logic Group Preprint Series 216, Department of Philosophy, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, October 2002.

[Vuk96] M Vuković. Some correspondence of principles in interpretability logic. *Glasink Matematički*, 31(51):193–200, 1996.