Modal Matters in Interpretability Logics

Evan Goris and
Joost J. Joosten

March 12, 2004

Abstract

In this paper we expose a method for building models for interpretability logics. The method can be compared to the method of taking unions of chains in classical model theory. Many applications of the method share a common part. We isolate this common part in a main lemma. Doing so, many of our results become applications of this main lemma. We also briefly describe how our method can be generalized to modal logics with a different signature.

With the general method, we prove completeness for the interpretability logics **IL**, **IL**M, **IL**M_0 and **IL**W^*. We also apply our method to obtain a classification of the essential Σ_1-sentences of essentially reflexive theories. We briefly comment on such a classification for finitely axiomatizable theories. As a digression we proof some results on self-provers.

Towards the end of the paper we concentrate on modal matters concerning **IL**(All), the interpretability logic of all reasonable arithmetical theories. We prove the modal incompleteness of the logic **IL**$\mathsf{W}^*\mathsf{P}_0$. We put forward a new principle R, and show it to be arithmetically sound in any reasonable arithmetical theory. Finally we make some general remarks on the logics **IL**RW and **IL**(All).

Contents

1 Introduction

In this paper we describe in great detail a method for modal completeness proofs. The method we present is actually not new. But, to the best of our knowledge, it is the first time it has been given a thorough and complete treatment of its own. The reason for starting such a study was twofold.

In the first place there is rigor. Modal completeness in interpretability logics is often a rather involved and elaborate business. Therefore, almost never a fully detailed completeness proof has been given. Mistakes are easily made, and indeed have been made in many papers. We thought it desirable to, at least once, provide a fully detailed completeness proof of a non-trivial interpretability logic.

In the second place, we hoped to provide a tool-kit that would make rigorous proofs easier and shorter. To a certain extend we think we have succeeded in doing so. We have tried to isolate a part of reasoning present in all completeness proofs for interpretability logics. That is to say, all completeness proofs that are presented in the format we have developed in this paper.

Developing the necessary toolkit is sometimes rather tough-going and tedious. We do think however that the results are somehow rewarding.

1.1 How to (not) read this paper.

Indeed, this paper is a rather lengthy one. We see two reasons for this excessive length. First of all, as we already mentioned, we have chosen to write down fully detailed proofs. This is merely for the sake of correctness. Not rarely, the word "trivial" in a modal completeness proof, is actually an indication of an error or small mistake. Of course, once the reader is convinced of some claim, all the tedious checks can be skipped, and we actually encourage the reader to do so.

Secondly, the paper deals with many different modal matters, as the title already indicates. Consequently, the paper gets some encyclopaedic flavor. Therefore one should not try to read the paper linearly but rather look up the result one is interested in.

2 Interpretability logics

In this section we introduce the basic notions that are used in the rest of this paper. We tried to keep the treatment self contained. It is advised to only turn to this section to look up a notion or a basic fact about 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 $\rightarrow$ and $\leftrightarrow$. We shall also omit outer brackets. Thus, we shall write $A \rhd B \rightarrow A \wedge \Box C \rhd B \wedge \Box C$ instead of $((A \rhd B) \rightarrow ((A \wedge (\Box C)) \rhd (B \wedge (\Box C))))$.

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 $A \rightarrow B$ 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 $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$
- L2 $\Box A \rightarrow \Box\Box A$
- L3 $\Box(\Box A \rightarrow A) \rightarrow \Box A$
- J1 $\Box(A \rightarrow B) \rightarrow A \rhd B$
- J2 $(A \rhd B) \wedge (B \rhd C) \rightarrow A \rhd C$
- J3 $(A \rhd C) \wedge (B \rhd C) \rightarrow A \vee B \rhd C$
- J4 $A \rhd B \rightarrow (\Diamond A \rightarrow \Diamond B)$
- 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 $\mathbf{IL} \vdash \varphi \rightarrow \psi \rightarrow \chi$ as short for $\mathbf{IL} \vdash \varphi \rightarrow \psi$ & $\mathbf{IL} \vdash \psi \rightarrow \chi$. 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 $\mathsf{L1}$-$\mathsf{L3}$. The axiom schema $\mathsf{L3}$ 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 $\mathsf{L3}$ 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 $\mathsf{J5}$ 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 $\rhd$. $\dashv$

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

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

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

$\mathsf{M_0}$ $A \rhd B \to \Diamond A \wedge \Box C \rhd B \wedge \Box C$

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

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

$\mathsf{P_0}$ $A \rhd \Diamond B \to \Box(A \rhd B)$

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

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 **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 **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. For, let $\sigma, A \to B$ be an $\mathbf{IL}\mathsf{X}$-proof of $A \to B$ from Γ. Then $\sigma, A \to B, A, B$ is an $\mathbf{IL}\mathsf{X}$-proof of B from Γ, A.

"$\Rightarrow$" goes by induction on the length n of the $\mathbf{IL}\mathsf{X}$-proof σ of B from Γ, A. If n=1, then $\sigma = B$ and $B \in \Gamma \cup \{A\}$. If $B = A$, clearly $\Gamma \vdash_{\mathbf{IL}\mathsf{X}} A \to A$. If $B \in \Gamma$, also $\Gamma \vdash_{\mathbf{IL}\mathsf{X}} B$.

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. Namely $\sigma', A \to C, \sigma'', A \to (C \to B), (A \to (C \to B)) \to ((A \to C) \to (A \to B)), (A \to C) \to (A \to B), A \to B$. ⊣

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. ⊣

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.

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* $\mathbf{IL}\mathsf{X}$*-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 $\neg A, \Box\neg A \in \Delta$.

⊣

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.

- $w\widetilde{\Vdash}p$ iff. $w \Vdash p$ for $p \in \mathsf{Prop}$
- $w \;\widetilde{\nVdash} \bot$
- $w\widetilde{\Vdash}A \rightarrow B$ iff. $w \;\widetilde{\nVdash} A$ 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. $\forall u\; (wRu \wedge u\widetilde{\Vdash}A \Rightarrow \exists v\; (uS_wv\widetilde{\Vdash}B))$

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 there is chance of confusion, 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.

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

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 will only comment on one direction of the case $\varphi = A \rhd B$. So, we suppose that for some $x \in m{\upharpoonright}*$ we have $M, x \Vdash A \rhd B$, and will show that $M{\upharpoonright}m, x \Vdash A \rhd B$. If now $M{\upharpoonright}m, y \Vdash A$ with $M{\upharpoonright}m \models xRy$, then also $M \models xRy$, whence, by the induction hypothesis, $M, y \models A$. We can thus find a z with $M \models yS_xz$ and $M, z \Vdash B$. As $x \in m{\upharpoonright}*$, we see that $M{\upharpoonright}m \models yS_xz$. By the induction hypothesis $M{\upharpoonright}m, z \Vdash B$. $\dashv$

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.

[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.

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.

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 **IL**X. A frame satisfying the **IL**X frame condition is often called an **IL**X-frame. In case no such frame condition exists, an **IL**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 **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* **IL**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. ⊣

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

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 $\mathbf{IL}\mathsf{X} \vdash \varphi \Rightarrow M \models \varphi$.*

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). ⊣

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, $\mathsf{Prov}_T(\varphi) := \exists p \; \mathsf{Proof}_T(p, \varphi)$.

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

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.

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

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 $\bot^* = (\vee_\varnothing)^* = \vee_\varnothing = \bot$.
- $(\Box\varphi)^* = \mathsf{Prov}_T(\varphi^*)$
- $(\varphi \rhd \psi)^* = Int_T(\varphi^*, \psi^*)$

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

In this section we will expose the main ingredients of our construction method. We will explain why we have chosen to work with these particular ingredients and compare them to other methods in the literature.

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 4.10 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

[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]. Boolos also gives an example of Goldfarb that uses only one propositional variable. Goldfarb uses infinitely many different formulas though. We note that conversely ill-foundedness can not be imposed by a consistent (over some complete logic $\mathbf{IL}\mathsf{X}$) set of finitely many formulas. This is because the construction method, or even just the completeness, would yield a conversely well-founded model forcing these finitely many formulas in some point.

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.

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 6.9 and Lemma 11.15.

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 4.12)

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

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 $x \not\Vdash \neg(\varphi \rhd \psi)$. We define a deficiency to be a configuration such that $\varphi \rhd \psi \in x \cap \mathcal{D}$ but $x \not\Vdash \varphi \rhd \psi$. 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 reoccur. 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 [Bus98]). However, for logics that contain $\mathbf{IL}\mathsf{M}_0$ 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 $\mathbf{IL}$ 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.[7] Alternatively, one could resort to other methods for showing decidability like the Mosaic method [BV01].

4 The construction method

In this section we will expose the construction method and prove its correctness. Most of the applications of the construction method are in proving some logic complete. In the end of this section we shall make some remarks on these sort of applications.

4.1 Preparing the construction; Some definitions

In this section we will make extensive use of maximal $\mathbf{IL}\mathsf{X}$-consistent sets for interpretability logics $\mathbf{IL}\mathsf{X}$. As we have seen in Section 3, in our setting the following definitions come very natural. Let us first recall Definition 2.7 and introduce some more nomenclature.

Definition 4.1. Let Γ and Δ be maximal $\mathbf{IL}\mathsf{X}$-consistent sets. We set $\Gamma \prec \Delta$ iff. $[\Box A \in \Gamma \Rightarrow A, \Box A \in \Delta]$. We say that Δ is a successor of Γ.

Definition 4.2. Let Γ and Δ be maximal $\mathbf{IL}\mathsf{X}$-consistent sets and let C be some formula. We set $\Gamma \prec_C \Delta$ iff. $[B \rhd C \in \Gamma \Rightarrow \neg B, \Box\neg B \in \Delta]$. We call Δ a C-critical successor of Γ.

It is clear that C-critical successors are indeed successors. For, let $\Box A \in \Gamma$. As $\mathbf{IL} \vdash \Box A \leftrightarrow \neg A \rhd \bot$ and Γ is maximal $\mathbf{IL}\mathsf{X}$-consistent, we have $\neg A \rhd \bot \in \Gamma$. Clearly $\bot \rhd C \in \Gamma$ and thus also $\neg A \rhd C \in \Gamma$. By C-criticallity we get $A, \Box A \in \Delta$.

Definition 4.3. Let Γ and Δ be maximal $\mathbf{IL}\mathsf{X}$-consistent sets. We set $\Gamma \subseteq_\Box \Delta$ iff. $[\Box A \in \Gamma \Rightarrow \Box A \in \Delta]$.

We now come to a central definition in our construction; that of an $\mathbf{IL}\mathsf{X}$-labeled frame. 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 4.4. An $\mathbf{IL}\mathsf{X}$-*labeled frame* is a quadruple $\langle W, R, S, \nu\rangle$. Here $\langle W, R, S\rangle$ is an $\mathbf{IL}$-frame and ν is a labeling function. The function ν

[7] In [Joo98] a sketch is given to obtain decidability for $\mathbf{IL}\mathsf{M}_0$.

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 upon 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 $xR^A y$ or $\nu(x, y) = A$.

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

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

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 $\mathsf{Form}_{\mathbf{IL}} \cup \{\Gamma \mid \Gamma$ is a maximal $\mathbf{IL}\mathsf{X}$ consistent set$\}$

Remark 4.5. Every $\mathbf{IL}\mathsf{X}$-labeled frame $F = \langle W, R, S, \nu\rangle$ can be transformed to an $\mathbf{IL}$-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 $\mathbf{IL}\mathsf{X}$ frame condition[8] as an $\mathbf{IL}\mathsf{X}$-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 $\mathbf{IL}\mathsf{X}$-models M, we have $\overline{\overline{M}} = M$, but in general $\overline{\overline{F}} \neq F$ for $\mathbf{IL}\mathsf{X}$-labeled frames F.

Definition 4.6. Let x be a world in some $\mathbf{IL}\mathsf{X}$-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 4.7. Let x be a world in some $\mathbf{IL}\mathsf{X}$-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 4.6 and 4.7. The last clause in the inductive definitions demands closure of the cone

[8] We could even say, any $\mathbf{IL}\mathsf{X}$-model.

under R-successors. But from Definition 2.9.5 closure of the cone under R follows from closure of the cone under S_x. And closure of the cone under S_x follows from the second clause. 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 $(x{\upharpoonright})^2 \cap R \subseteq S_x$.

In our construction we will gradually build up **IL**X-labeled frames so that the corresponding models get more and more of the desired properties. The following two definitions provide a means to measure how many of our desiderata we still miss.

Definition 4.8. [9] 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 $[A \in \nu(y) \Rightarrow \exists z\ (yS_xz \wedge B \in \nu(z))]$.

Definition 4.9. 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_xz we have $D \in \nu(z)$.

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

Definition 4.10. 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".

Definition 4.11. 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 4.12. *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. We have included it though, to show the interplay between all the ingredients.

Suppose a truth lemma holds. To see that there are no $\mathcal{D}$-problems we consider some $\neg(A \rhd B) \in \mathcal{D}$ and some $x \in F$ with $\neg(A \rhd B) \in \nu(x)$. We need to find a y with xRy, $A \in \nu(y)$ such that for no z with yS_xz, $B \in \nu(z)$. As a truth lemma holds and $\neg(A \rhd B) \in \nu(x)$, we have that $\overline{F}, x \Vdash \neg(A \rhd B)$, and thus, for some y we have: $xRy \Vdash A$ and for no z with yS_xz, holds $\overline{F}, z \Vdash B$. Thus, $yS_xz \rightarrow z \Vdash \neg B$. As $\mathcal{D}$ is closed

[9] This is a temporary definition. We will eventually work with Definition 4.14.

under subformulas, $B \in \mathcal{D}$. And, as $\mathcal{D}$ is closed under single negations, $\sim B \in \mathcal{D}$. Clearly $\mathbf{IL} \vdash \neg B \leftrightarrow \sim B$ and thus, $yS_xz \rightarrow z \Vdash \sim B$. By the truth lemma, we have that for any z, $yS_xz \rightarrow \sim B \in \nu(z)$. As $\nu(z)$ is a maximal consistent set, $\sim B \in \nu(z) \Rightarrow \neg B \in \nu(z)$ and thus we see that $\langle x, \neg(A \rhd B)\rangle$ is not a $\mathcal{D}$-problem.

To see that there are no deficiencies, we reason as follows. Consider xRy with $C \rhd D \in \nu(x) \cap \mathcal{D}$, and $C \in \nu(y)$. We need to find a z with yS_xz and $D \in \nu(z)$. Again, $C \rhd D \in \nu(x)$ implies by the truth lemma that $\overline{F}, x \Vdash C \rhd D$. Analogously we get $\overline{F}, y \Vdash C$. Thus, we find z with $yS_xz \Vdash D$. Again, by the truth lemma and the closure of $\mathcal{D}$ under subformulas, we see that $D \in \nu(z)$. Thus $\langle x, y, C \rhd D\rangle$ can not be a deficiency.

For the other direction, assume that there are no $\mathcal{D}$-problems nor $\mathcal{D}$-deficiencies. We will prove by induction on the complexity of $A \in \mathcal{D}$ that

$$\forall x \, (\overline{F}, x \Vdash A \text{ iff. } A \in \nu(x)).$$

For propositional variables this holds by our definition of $\Vdash$. The logical structure of maximal consistent sets mimics very well the behavior of $\Vdash$ when it comes to boolean connectives. For example $\overline{F}, x \Vdash A \wedge B \Leftrightarrow \overline{F}, x \Vdash A \;\&\; \overline{F}, x \Vdash B \Leftrightarrow_{\text{I.H.}} A \in \nu(x) \;\&\; B \in \nu(x) \Leftrightarrow A \wedge B \in \nu(x)$. Note that we have access to the induction hypothesis as $\mathcal{D}$ is closed under subformulas. For the negation we see that $\overline{F}, x \Vdash \neg A \Leftrightarrow \overline{F}, x \nVdash A \Leftrightarrow_{\text{I.H.}} A \notin \nu(x) \Leftrightarrow \neg A \in \nu(x)$. Note that in this case we have access to the induction hypothesis again by the subformula property.

We now need to see that $\overline{F}x \Vdash C \rhd D \Leftrightarrow C \rhd D \in \nu(x)$ for any $C \rhd D \in \mathcal{D}$. This part is split in two directions.

First suppose that $C \rhd D \in \nu(x)$. We need to see that $x \Vdash C \rhd D$. Thus, suppose that $xRy \Vdash C$. By the induction hypothesis $C \in \nu(y)$ and, as there are no deficiencies, we can find z with $yS_xz \ni D$. Again by the induction hypothesis, $z \Vdash D$ and indeed $x \Vdash C \rhd D$.

For the other direction suppose that $C \rhd D \notin \nu(x)$. Thus $\neg(C \rhd D) \in \nu(x)$. Note that $\sim(C \rhd D) = \neg(C \rhd D)$. As there are no $\mathcal{D}$-problems and $\neg(C \rhd D) \in \mathcal{D}$, we can find some y with xRy and $C \in \nu(y)$ and for no z with yS_xz, $D \in \nu(z)$. Or equivalently, for all z with yS_xz, $D \notin \nu(z)$. By applying twice the induction hypothesis we see that $x \Vdash \neg(C \rhd D)$. $\dashv$

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

Definition 4.13 (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 4.14 (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 $A \in \nu(y)$.

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 4.15. It is also easy to see that the we still have the interesting half of Lemma 4.12. 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 4.16 (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'_xz iff. yS_xz
- $\nu'(x) = \nu(x)$
- $\nu'(\langle x, y\rangle)$ 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 4.17 (Depth). The *depth* of a finite frame F, we will write $\mathsf{depth}(F)$ is the maximal length of sequences of the form $x_0R\ldots Rx_n$. (For convenience we define $\max(\varnothing) = 0$.)

Definition 4.18 (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 the 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 4.19 (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:* $xRy \rightarrow \exists A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box\neg D \mid D \text{ a subformula of some } B \in \mathcal{D}\}$.
- *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.*

It is clear that the lemma will be proved by subsequently eliminating problems and deficiencies by means of extensions. These elimination processes have to be robust in a sense that every problem or deficiency that has been dealt with, should not possibly re-emerge. But, as we shall see, the requirements of the lemma almost immediately imply this. The following two lemmata however show that the requirements are not so strong. The first lemma relates to problems.

Lemma 4.20. *Let* Γ *be a maximal* $\mathbf{IL}\mathsf{X}$*-consistent set such that* $\neg(A \rhd B) \in \Gamma$*. Then there exists a maximal* $\mathbf{IL}\mathsf{X}$*-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[10] formula C for which $C \rhd B \in \Gamma$ such that

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

[10]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 $C \rhd B \in \Gamma$.

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. $\dashv$

For deficiencies there is a similar lemma.

Lemma 4.21. *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 4.20 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 $\Gamma \prec_B \Delta \ni C$. $\dashv$

In the proof of our main lemma however we will not consider the process of eliminating problems and deficiencies in that much detail. We required a sort of black box that does the eliminations in the conditions of the theorem. Every application however will make use of Lemmas 4.20 and 4.21.

Proof of Lemma 4.19. So, let $\mathbf{IL}\mathsf{X}$, $\mathcal{D}$, $\mathcal{C}$ and $\mathcal{I}$ be given so that the requirements of the lemma are satisfied. We first proof that every labeled adequate frame F satisfying all the invariants can be extended to a labeled adequate $\mathbf{IL}\mathsf{X}$-frame $\hat{F}$ on which a truth lemma w.r.t. $\mathcal{D}$ holds.

In the light of Lemma 4.12 and of Remark 4.15 we are done if we can find an extension of F where no $\mathcal{D}$-problems nor $\mathcal{D}$-deficiencies occur.

Actually, we will assume that $\mathcal{D}$ is closed under subformulas and single negations. If $\mathcal{D}$ does not have these closure properties, we can first close $\mathcal{D}$ off to get a set $\mathcal{D}'$ that does have the closure properties. Clearly $\mathcal{D}'$ is also a finite set. Thus, without loss of generality we may assume that $\mathcal{D}$ is closed under subformulas and single negations. In this case $\{\Box\neg D \mid D \text{ a subformula of some } B \in \mathcal{D}\} = \{\Box\neg D \mid D \in \mathcal{D}\} = \{\Box D \mid D \in \mathcal{D}\}$, where the last equality is not really an equality but rather some sort of equivalence.

The idea of the proof is very simple. We start with $F_0 := F$ and consider some deficiency or problem in it. We eliminate this problem or deficiency by extending F_0 to F_1. Next we consider some problem or deficiency in F_1 and eliminate it by extending F_1 to F_2. Proceeding like this we get a (possibly) infinite chain.

$$F = F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots \subseteq \cup_{i\in\omega} F_i =: \hat{F} \quad (i)$$

As we shall see, this $\hat{F}$ will be our required extension of F if we choose our intermediate F_i right. At this moment we can point out four points of attention.

1. Newly created problems and deficiencies should also at some point be eliminated.

2. Problems and deficiencies that have been eliminated, should not come back at a later stage.

3. The chain (i) should be a bounded chain.

4. The limit should be an adequate labeled **IL**X-frame containing no problems and no deficiencies.

We now see how these points get incorporated in the construction.

Point 1 is really not problematic. We can just take care of it by fixing some enumeration of problems and deficiencies. To this extend, we fix a countable infinite set of names $X := \{x_0, x_1, \dots\}$ for our current and future worlds. Every world in some F_i will be some $x \in X$. Next we consider the set $\mathcal{A} := \{\langle x, \neg(A \rhd B)\rangle \mid x \in X, \neg(A \rhd B) \in \mathcal{D}\} \cup \{\langle x, y, C \rhd D\rangle \mid x, y \in X, C \rhd D \in \mathcal{D}\}$ and we fix some enumeration on $\mathcal{A}$. If we are to choose at a certain stage some deficiency or problem to eliminate, we just pick the least (with respect to the enumeration order) element of $\mathcal{A}$ that is indeed a problem or a deficiency. If we now know that problems and deficiencies, once dealt with, will never re-occur, we are sure that we come higher and higher in the enumeration of $\mathcal{A}$. Point 2 precisely deals with the robustness of the elimination method.

Point 2. It is easy to see that deficiencies, once eliminated by means of an extension, will never re-occur. Consider $C \rhd D \in \nu(x)$ and $C \in \nu(y)$ and xRy. If $\langle x, y, C \rhd D\rangle$ is a deficiency in F_i that is eliminated at this stage, it will be eliminated by adding (at least) a new element z to F_i. Thus, F_{i+1} will contain z with $D \in \nu(z)$ and yS_xz. This world z will also be in all extensions of F_{i+1}.

To see that we can eliminate problems in such a way, so that they will never re-occur, we have to be a bit more precise. Let $\langle x, \neg(A \rhd B)\rangle$ be a problem of F_i that will be eliminated in F_{i+1}. Thus, some $y \in \mathcal{C}_x^B$ is added with $A \in \nu(y)$. We need to see that in no F_j, $j \geq i+1$ there is a z with yS_xz and $B \in \nu(z)$. But if yS_xz, we have by the definition of $\mathcal{C}_x^B$ that $z \in \mathcal{C}_x^B$. By adequacy we see that[11] $x \prec_B z$ and thus $\neg B \in \nu(z)$.

Point 3. We should provide a bound on $\mathsf{depth}(F_i)$ of the elements of our chain (i). This is taken care of by the invariant $xRy \rightarrow \exists\, A {\in} (\nu(y) \setminus \nu(x)) \cap \{\Box\neg D \mid D \text{ a subformula of some } B \in \mathcal{D}\}$. Clearly, if in some F_i we have that $x_0Rx_1R\dots Rx_m$ we have that $m \leq |\mathcal{D}|$.

Point 4. We should have that $\hat{F} := \cup_{i\in\omega} F_i$ is a labeled adequate **IL**X-frame. For adequacy we should check a list of items. Amongst these are: transitivity of R, conversely well-foundedness of R, reflexivity and transitivity of S_x, $xRyRz \rightarrow yS_xz$, $yS_xz \rightarrow xRz$. It is completely straightforward to show that these properties are preserved under taking bounded unions of chains. As $\hat{F} \models \mathcal{I}^{\mathcal{U}}$, we get from our assumption that $\hat{F} \models \mathcal{C}$ and thus $\hat{F}$ is an **IL**X-frame. Clearly $\hat{F}$ can not have any problems or deficiencies and thus a truth lemma holds in $\hat{F}$ with respect to $\mathcal{D}$.

This proves the first part of the Main Lemma.

We will now prove the second part of the Main Lemma. Thus, 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

[11] This is actually the only property of adequacy that is used in the proof of the main lemma.

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 $\hat{F}$. ⊣

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.

4.3 How to use 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 4.22. A modal completeness proof for an interpretability logic $\mathbf{IL}\mathsf{X}$ 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 $xRy \to \exists\, A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}$.
- **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 $\mathbf{IL}$-model. For example, in general, the R relation is not transitive. To come back to at least an $\mathbf{IL}$-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 $\mathbf{IL}\mathsf{X}$ and a corresponding notion of an $\mathbf{IL}\mathsf{X}$-frame. (Cf. Definition 2.9 and Lemma 2.15 in the case of **IL**, Definition 6.1 and Lemma 6.2 in the case of $\mathbf{IL}\mathsf{M}$, Theorem 11.18 in the case of $\mathbf{IL}\mathsf{M}_0$ and Theorem 12.11 in the case of $\mathbf{IL}\mathsf{W}^*$.)
- Defining a notion of a quasi $\mathbf{IL}\mathsf{X}$-frame. (Cf. Definition 5.1 in the case of **IL**, Definition 6.4 in the case of $\mathbf{IL}\mathsf{M}$, Definition 11.9 in the case of $\mathbf{IL}\mathsf{M}_0$ and Definition 12.6 in the case of $\mathbf{IL}\mathsf{W}^*$.)
- Defining some notions that remain constant throughout the closing of quasi $\mathbf{IL}\mathsf{X}$-frames, but somehow capture the dynamic features of this process. (Cf. critical and generalized cones; Definitions 4.6 and 4.7, in the case of **IL**; critical $\mathcal{M}$-cone, Definition 6.3, and $R \circ S$ in the case of $\mathbf{IL}\mathsf{M}$ and Definition 11.5 and Definition 11.8 both in the case of $\mathbf{IL}\mathsf{M}_0$ and $\mathbf{IL}\mathsf{W}^*$.)
- Proving that a quasi $\mathbf{IL}\mathsf{X}$-frame can be closed off to an adequate labeled $\mathbf{IL}\mathsf{X}$-frame. (Cf. Lemma 5.2 and corollary 5.3 for **IL**, Lemma 6.5 and Corollary 6.8 for $\mathbf{IL}\mathsf{M}$, Lemma 11.12 in the case of $\mathbf{IL}\mathsf{M}_0$ and Lemma 12.8 in the case of $\mathbf{IL}\mathsf{W}^*$.)
- Preparing the elimination of deficiencies. (Cf. Lemma 4.21 in the case of **IL**, 6.9 in the case of $\mathbf{IL}\mathsf{M}$, Lemma 11.15 in the case of $\mathbf{IL}\mathsf{M}_0$ and Lemma 12.9 in the case of $\mathbf{IL}\mathsf{W}^*$.)

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 $\mathbf{IL}\mathsf{X}$-frames to adequate $\mathbf{IL}\mathsf{X}$-frames is a very laborious enterprise. We have chosen to do it in great detail though as it is the place where all the essential ingredients come together. Furthermore, this work can be used time and again, once it is executed.

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. (Enumerations are always of type ω.) 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) $F_{n+1} := \langle W_n, R_n \cup \{\langle a, c \rangle\}, S_n \cup \{\langle a, b, c \rangle\}, \nu_n \rangle$

We first see by induction on n that all the F_n are quasi-frames. For $n = 0$ this is true by definition.

We follow the case distinctions and see that at every step we obtain a quasi-frame. In order to see that we have a quasi-frame, the following five requirements should be checked.

1. R is conversely well-founded
2. $yS_x z \rightarrow xRy$ & 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. $y \in \mathcal{C}_x^A \rightarrow \nu(x) \prec_A \nu(y)$

Instead of proving 4 and 5 we will prove two stronger statements. We will prove that $\mathcal{G}_x^A$ and $\mathcal{C}_x^A$ are actually the same set for all n. Clearly 4 and 5 will follow from this observation and the fact that F_0 is a quasi-frame. What we would like to prove can be expressed by the following.

4". $\forall n \forall m\ [F_n \models y \in \mathcal{G}_x^A \Leftrightarrow F_m \models y \in \mathcal{G}_x^A]$

5". $\forall n \forall m\ [F_n \models y \in \mathcal{C}_x^A \Leftrightarrow F_m \models y \in \mathcal{C}_x^A]$

For this, it is sufficient to prove that.

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

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

We can now just do all the cases. In case $F_{n+1} = F_n$ we are immediately done by the induction hypothesis.

Case (i): We have eliminated an imperfection concerning the transitivity of the R relation and $F_{n+1} := \langle W_n, R_n \cup \{\langle a, c\rangle\}, S_n, \nu_n\rangle$. In this case, 2 is clear as $S_{n+1} = S_n$. For 1, suppose a sequence $x_0, x_1, \ldots$ is given such that $F_{n+1} \models x_0 R x_1 R \ldots$. This sequence $x_0, x_1, \ldots$ can be transformed into a sequence $y_0, y_1, \ldots$ such that $F_n \models y_0 R y_1 R \ldots$ by replacing all occurrences of aRc in $x_0, x_1, \ldots$ by $aRbRc$ and leaving the rest unchanged. Clearly, if $x_0, x_1, \ldots$ is infinite, also $y_0, y_1, \ldots$ is infinite. We conclude that $x_0, x_1, \ldots$ can not be infinite and 1 holds on F_{n+1}. To see 3 we only need to check that $\nu(a) \prec \nu(c)$. This follows from the transitivity of the $\prec$ relation and $\nu(a) \prec \nu(b) \prec \nu(c)$. To see 4' we reason as follows. Suppose $F_{n+1} \models y \in \mathcal{G}_x^A$. By the inductive definition this means that there are $x_0, \ldots, x_k$ such that $F_{n+1} \models xR^A x_0 Q x_1 Q \ldots Q x_k Q y$, where $Q \in \{R, S\}$. Recall that uSv means that $uS_w v$ for some w. We get a sequence $y_0, \ldots, y_m$ such that $F_n \models xR^A y_0 Q \ldots Q y_m Q y$ by replacing all occurrences of aRc by $aRbRc$ in $x_0, \ldots, x_k$ and leaving the rest unchanged. Thus $F_{n+1} \models y \in \mathcal{G}_x^A \Rightarrow F_n \models y \in \mathcal{G}_x^A$. The other direction is obviously true. The validity of 5' follows from a completely analogous reasoning.

Case (ii): We have eliminated an imperfection concerning the reflexivity of S_a and $F_{n+1} := \langle W_n, R_n, S_n \cup \{\langle a, b, b\rangle\}, \nu_n\rangle$. In this case, 1 and 3 are clear as $R_n = R_{n+1}$. Furthermore, 2 is also clear, as we have that $\langle a, b\rangle \in R_n$. As before we see that 4' and 5' hold.

Case (iii): We have eliminated an imperfection concerning the transitivity of S_a and $F_{n+1} := \langle W_n, R_n, S_n \cup \{\langle a, b, d\rangle\}, \nu_n\rangle$. Again 1 and 3 are clear. The assumption tells us that bS_acS_ad. Thus by the induction hypothesis, aRd whence 2 is also satisfied. As before we see that 4' and 5' hold.

Case (iv): We have eliminated an imperfection concerning the inclusion of R in S_a and $F_{n+1} := \langle W_n, R_n \cup \{\langle a, c\rangle\}, S_n \cup \{\langle a, b, c\rangle\}, \nu_n\rangle$. To see that 1 and 3 hold we apply precisely the same argument as in Case (i). The only difference now is that $F_n \models aRc$ is also possible, which would even simplify the argument. To see that 2 holds, we only need to consider the newly added bS_ac. But $R_{n+1} = R_n \cup \{\langle a, c\rangle\}$, thus certainly $F_{n+1} \models aRc$. Points 4' and 5' go as always.

Thus indeed F_n is a quasi-frame for every n.

We will now see that $F := \cup_{i\in\omega} F_i$ is the required adequate **IL**-frame. To this extend we have to establish the following properties.

(a.) W is the domain of F

(b.) $R_0 \subseteq \cup_{i\in\omega} R_i$

(c.) $S_0 \subseteq \cup_{i\in\omega} S_i$

(d.) R is conversely well-founded on F

(e.) $F \models xRyRz \rightarrow xRz$

(f.) $F \models yS_xz \rightarrow xRy\ \&\ xRz$

(g.) $F \models xRy \rightarrow yS_xy$

(h.) $F \models xRyRz \rightarrow yS_xz$

(i.) $F \models uS_xvS_xw \rightarrow uS_xw$

(j.) $F \models xRy \Rightarrow \nu(x) \prec \nu(y)$

(k.) $A \neq B \Rightarrow F \models \mathcal{G}^A_x \cap \mathcal{G}^B_x = \varnothing$

(l.) $F \models y \Subset \mathcal{C}^A_x \Rightarrow \nu(x) \prec_A \nu(y)$

Properties ($a.$)-($c.$) say that F is an extension of G. Properties ($d.$)-($i.$) are from Definition 2.9 and Properties ($j.$)-($l.$) are the adequateness conditions.

Properties ($a.$), ($b.$) and ($c.$) are obvious. To see ($d.$) we reason as follows. We first prove by induction on n that if $F_n \models xRy$, then there are $a_1, \ldots, a_m (0 \leq m)$ such that $F_0 \models xRa_1R\ldots Ra_mRy$. By this property any chain $x_0, x_1, \ldots$ with $F \models x_0Rx_1R\ldots$ can be transformed into a chain $y_0, y_1 \ldots$ with $F_0 \models y_0Ry_1R\ldots$. Clearly, if $x_0, x_1, \ldots$ is infinite, then so is $y_0, y_1 \ldots$. Thus by the conversely well-foundedness of F_0 we see that F is conversely well-founded.

To see ($f.$), suppose $F \models yS_xz$. Then, for some n, $F_n \models yS_xz$. As F_n is a quasi-frame, $F_n \models xRy$ and $F_n \models xRz$. Thus the same holds in F. Property ($j.$) is proved similarly.

The validity of ($k.$) and ($l.$) is immediate from our previous observations 4" and 5" that the $\mathcal{G}^A_x$ and the $\mathcal{C}^A_x$ entities are stable throughout the chain.

For the remaining cases ($e.$), ($g.$), ($h.$) and ($i.$) the following two observations on imperfections are central.

- If γ is not an imperfection in F_n, then it will not be an imperfection in F_m for any $m \geq n$. So, certainly it will not be an imperfection of F.
- If γ is an imperfection of F_n, then there are only finitely many imperfections $\gamma' <_0 \gamma$ in F_n.

We now see in one go that no imperfections can hold in F. Assume for a contradiction that γ is an imperfection in F. Thus, for some n, γ is an imperfection of F_n. There are only finitely many, say m, imperfections γ' below γ w.r.t. the $<_0$ ordering. At each stage one of these γ' will disappear. Thus γ is not an imperfection in F_{n+m+1} and hence not in F. A contradiction. $\dashv$

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 4.21 suffices.

5.2 Modal completeness

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

Proof. We specify the four ingredients mentioned in Remark 4.22.

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 4.20 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 $b \in \mathcal{C}^C_x$ for $x{\neq}a$, that then $\nu(x) \prec_C \nu(b)$. For, $b \in \mathcal{C}^C_x$ implies $a \in \mathcal{C}^C_x$ 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 $b \in \mathcal{C}^B_a$. 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 $b \in \mathcal{C}^B_a$. 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 4.21 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 **IL**M 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 (see Definition 6.1). Thus, 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.

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

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. $F \models \forall x, y, u, v\ (yS_xuRv \to yRv) \Leftrightarrow F \models$ **IL**M

Proof. "$\Rightarrow$". We take any model M based on F and $a \in M$. We suppose $a \Vdash A \rhd B$. Thus, if $aRb \Vdash A \land \Box C$, then there is some c with $aRbS_ac \Vdash B$. If cRd, by the frame condition bRd whence $d \Vdash C$. Thus $c \Vdash B \land \Box C$ and $a \Vdash A \land \Box C \rhd B \land \Box C$.

"$\Leftarrow$". We consider $a, b, c, d \in F$ with bS_acRd. We define a model M, based on F, by specifying $\Vdash$.

$$\begin{array}{lcl} x \Vdash p & \Leftrightarrow & x = b \\ x \Vdash q & \Leftrightarrow & x = c \\ x \Vdash r & \Leftrightarrow & bRx \end{array}$$

Thus, $a \Vdash p \rhd q$, whence $a \Vdash p \wedge \Box r \rhd q \wedge \Box r$. Consequently $c \Vdash q \wedge \Box r$, which can only be the case if $cRx \to bRx$ for all x. Thus certainly bRd.

$\dashv$

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 $\mathcal{M}_x^A$, is defined inductively as follows.

- $xR^A y \to y \in \mathcal{M}_x^A$
- $y \in \mathcal{M}_x^A$ & $yRz \to z \in \mathcal{M}_x^A$
- $y \in \mathcal{M}_x^A$ & $yS_x z \to z \in \mathcal{M}_x^A$
- $y \in \mathcal{M}_x^A$ & $yS^{\mathsf{tr}}uRv \to v \in \mathcal{M}_x^A$

Definition 6.4. A quasi-frame is a quasi-**IL**M-frame if[12] the following properties hold.

- $R^{\mathsf{tr}}; S^{\mathsf{tr}}$ is conversely well-founded[13]
- $yS_x z \to \nu(y) \subseteq_\Box \nu(z)$
- $y \in \mathcal{M}_x^A \Rightarrow \nu(x) \prec_A \nu(y)$

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 having one of the following forms.

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

[12] 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.

[13] 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.

(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 $F_n \not\models bS_ad$

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

(v) $\gamma = \langle 4, a, b, c, d\rangle$ with $F_n \models bS_acRd$ but $F_n \not\models bRd$

We will define a chain of quasi-**IL**M-frames. Each new frame in the chain will have at least one imperfection less than its predecessor.

Let $<_0$ be the well-ordering on

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

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

$F_0 := G$

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

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

We will now see by induction on n that for all n, F_n is a quasi-**IL**M-frame. Thus we should check for the following list of properties.

1. R is conversely well-founded
2. $yS_xz \to xRy$ & xRz
3. $xRy \to \nu(x) \prec \nu(y)$
4. $A \neq B \to \mathcal{G}_x^A \cap \mathcal{G}_x^B = \varnothing$
5. $y \in \mathcal{C}_x^A \to \nu(x) \prec_A \nu(y)$
6. $R^{\mathsf{tr}}; S^{\mathsf{tr}}$ is conversely well-founded.
7. $yS_xz \to \nu(y) \subseteq_\Box \nu(z)$.
8. $y \in \mathcal{M}_x^A \to \nu(x) \prec_A \nu(y)$.

Again we follow our case distinction. We will only include the required new parts in the proof. All cases we do not deal with here, are dealt with in the proof of Lemma 5.2. Note that most of the proof of Lemma 5.2 can indeed be copied. There is just one exception. In Lemma 5.2 we prove that $\mathcal{C}_x^A$ is a constant entity, that is, the same set on all F_n. This is no longer the case in our current chain. Thus, Property 5' can no longer be proved. Instead we can prove that

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

that is, $\mathcal{M}_x^A$ is a constant entity. Clearly 8 follows from $8'$ as F_0 is a quasi-**IL**M-frame. As $\mathcal{C}_x^A \subseteq \mathcal{M}_x^A$, we see that 5 follows from 8.

Case (i): We have eliminated an imperfection concerning the transitivity of the R relation and $F_{n+1} := \langle W_n, R_n \cup \{\langle a, c\rangle\}, S_n, \nu_n\rangle$. Requirement 7 is obvious as $S_{n+1} = S_n$. To see 6, we reason as follows. We will see that $F_{n+1} \models xR^{\mathsf{tr}}y$ iff. $F_n \models xR^{\mathsf{tr}}y$. This is sufficient as $S_{n+1} = S_n$. It is clear that $F_n \models xR^{\mathsf{tr}}y \Rightarrow F_{n+1} \models xR^{\mathsf{tr}}y$. On the other hand, $F_{n+1} \models xR^{\mathsf{tr}}y$ iff. $\exists x_1, \ldots, x_m$ $(0 \leq m)$ $F_{n+1} \models xRx_1R \ldots Rx_mRy$. We go from $x_1, \ldots, x_m$ to $y_1, \ldots, y_l$ $(0 \leq l)$ by replacing every occurrence of aRc by $aRbRc$ and leaving the rest unchanged. Thus, $F_n \models xRy_1R \ldots Ry_lRy$ whence $F_n \models xR^{\mathsf{tr}}y$.

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[14] $F_{n+1} \models xR^A z_1(S_x \cup R \cup (S^{\mathsf{tr}}; R))z_2, \ldots, z_l(S_x \cup R \cup (S^{\mathsf{tr}}; R))y$. We transform the sequence $z_1, \ldots, z_l$ into a sequence $u_1, \ldots, u_m$ $(0 \leq m)$ in the following way. Every occurrence of aRc in $z_1, \ldots, z_l$ is replaced by $aRbRc$. In case that for some $n < l$ we have $z_nS^{\mathsf{tr}}aRc = z_{n+1}$, 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 $z_1, \ldots, z_l$ unchanged. Clearly $F_n \models xR^A u_1(S_x \cup R \cup (S^{\mathsf{tr}}; R))u_2, \ldots, u_m(S_x \cup R \cup (S^{\mathsf{tr}}; R))y$, whence $F_n \models y \in \mathcal{M}_x^A$.

Case (ii): We have eliminated an imperfection concerning the reflexivity of S_a and $F_{n+1} := \langle W_n, R_n, S_n \cup \{\langle a, b, b\rangle\}, \nu_n\rangle$. Now, 7 is obvious as $\nu(b) \subseteq_\Box \nu(b)$.

To see 6, we reason as follows. Suppose for a contradiction that we had an infinite sequence $x_1, x_2, x_3, \ldots$ such that $F_{n+1} \models x_1R^{\mathsf{tr}}x_2S^{\mathsf{tr}}x_3R^{\mathsf{tr}}x_4S^{\mathsf{tr}}x_5 \ldots$. We transform $x_1, x_2, x_3, \ldots$ into a sequence $y_1, y_2, y_3, \ldots$ as follows. If $x_{2m} = x_{2m+1} = b$ we just omit x_{2m+1}. In all other cases we do nothing.

Clearly, the thus obtained $y_1, y_2, y_3, \ldots$ is an infinite sequence too, as we only deleted elements with an odd index. We have for all i that $F_n \models y_iR^{\mathsf{tr}}y_{i+1}$ or $F_n \models y_iS^{\mathsf{tr}}y_{i+1}$. Moreover, for all i we have $F_n \models y_iR^{\mathsf{tr}}y_{i+1} \vee y_{i+1}R^{\mathsf{tr}}y_{i+2}$.

We consider two possible situations. One possible situation is that for all k from some j onwards, $F_n \models y_kR^{\mathsf{tr}}y_{k+1}$. In this situation R^{tr} would not be conversely well-founded on F_n. This is in contradiction with the assumption that R is conversely well-founded on F_n. In the other situation we have in F_n, infinitely many S^{tr}-transitions in our sequence $y_1, y_2, y_3, \ldots$. But this contradicts the assumption that $R^{\mathsf{tr}}; S^{\mathsf{tr}}$ is conversely well-founded on F_n. (We possibly use here that R^{tr} is a transitive relation.)

A similar reasoning shows us that $8'$ holds. If $F_{n+1} \models y \in \mathcal{M}_x^A$ then $\exists z_1, \ldots, z_l$ $(0 \leq l)$ with $F_{n+1} \models xR^A z_1(S_x \cup R \cup (S^{\mathsf{tr}}; R))z_2, \ldots, z_l(S_x \cup R \cup (S^{\mathsf{tr}}; R))y$. The only difference between F_{n+1} and F_n is that $F_{n+1} \models bS_ab$ and $F_n \not\models bS_ab$. This only has repercussions on S_a and S^{tr}. But as always, we can change to a sequence $u_1, \ldots, u_m$ $(0 \leq m)$ such that $F_n \models xR^A u_1(S_x \cup R \cup (S^{\mathsf{tr}}; R))u_2, \ldots, u_m(S_x \cup R \cup (S^{\mathsf{tr}}; R))y$, whence $F_n \models y \in \mathcal{M}_x^A$.

Case (iii): We have eliminated an imperfection concerning the transitivity of S_a and $F_{n+1} := \langle W_n, R_n, S_n \cup \{\langle a, b, d\rangle\}, \nu_n\rangle$. Again, 7 is easy,

[14]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. Clearly, the union operator is commutative and associative.

as $\subseteq_\Box$ is clearly a transitive relation on MCS's. In this case, to see 6, it suffices to remark that $F_{n+1} \models xS^{\mathsf{tr}}y$ iff. $F_n \models xS^{\mathsf{tr}}y$. The argument is similar to the one we exposed in Case (i).

To see $8'$ we reason as usual. Thus we transform a sequence such that $F_{n+1} \models xR^A z_1(S_x \cup R \cup (S^{\mathsf{tr}};R))z_2, \ldots, z_l(S_x \cup R \cup (S^{\mathsf{tr}};R))y$, $(0 \leq l)$, into a sequence such that $F_n \models xR^A u_1(S_x \cup R \cup (S^{\mathsf{tr}};R))u_2, \ldots, u_m(S_x \cup R \cup (S^{\mathsf{tr}};R))y$ $(0 \leq m)$. Clearly, S^{tr} does not change by adding $bS_a d$ to F_n. Thus it is obvious how to get from $z_1, \ldots, z_l$ to $u_1, \ldots, u_m$.

Case (iv): We have eliminated an imperfection concerning the inclusion of R in S_a and $F_{n+1} := \langle W_n, R_n \cup \{\langle a, c\rangle\}, S_n \cup \{\langle a, b, c\rangle\}, \nu_n\rangle$. Requirement 7 is easy, as $\nu(b) \prec \nu(c) \Rightarrow \nu(b) \subseteq_\Box \nu(c)$. To see 6, we reason as follows. As in Case (i) we see that $F_{n+1} \models xR^{\mathsf{tr}}y$ iff. $F_n \models xR^{\mathsf{tr}}y$. Now suppose for a contradiction that we had an infinite sequence $x_1, x_2, x_3, \ldots$ such that $F_{n+1} \models x_1 R^{\mathsf{tr}} x_2 S^{\mathsf{tr}} x_3 R^{\mathsf{tr}} x_4 S^{\mathsf{tr}} x_5 R^{\mathsf{tr}} x_6 S^{\mathsf{tr}} \ldots$. We will obtain an infinite sequence $y_1, y_2, y_3, \ldots$ such that[15]

$$F_n \models y_1 R^{\mathsf{tr}} y_2 S^{\mathsf{tr}} y_3 R^{\mathsf{tr}} y_4 \ldots \quad (*)$$

The new sequence will be the same as the old one apart from places where $F_{n+1} \models x_{2i} S^{\mathsf{tr}} x_{2i+1}$ (with $i > 0$) but $F_n \models \neg(x_{2i} S^{\mathsf{tr}} x_{2i+1})$. In this case either $x_{2i} = b$ and $x_{2i+1} = c$, or $x_{2i} = bRcS^{\mathsf{tr}} x_{2i+1}$, or $F_n \models x_{2i} S^{\mathsf{tr}} bRcS^{\mathsf{tr}} x_{2i+1}$, or $F_n \models x_{2i} S^{\mathsf{tr}} bRx_{2i+1} = c$. In all four cases it is clear how to proceed in order to obtain $(*)$. Clearly, also the $y_1, y_2, y_3, \ldots$ form an infinite sequence. If there are no infinite number of S^{tr}-transitions in the sequence, we get a contradiction with the fact that R is conversely well-founded on F_n. In the other case we contradict the assumption that $R^{\mathsf{tr}};S^{\mathsf{tr}}$ is a conversely well-founded relation on F_n.

To see $8'$ we reason as usual. Thus, we suppose $F_{n+1} \models y \in \mathcal{M}_x^A$ and we show that $F_n \models y \in \mathcal{M}_x^A$. We do this by transforming a sequence such that $F_{n+1} \models xR^A z_1(S_x \cup R \cup (S^{\mathsf{tr}};R))z_2, \ldots, z_l(S_x \cup R \cup (S^{\mathsf{tr}};R))y$ $(0 \leq l)$, into a sequence such that $F_n \models xR^A u_1(S_x \cup R \cup (S^{\mathsf{tr}};R))u_2, \ldots, u_m(S_x \cup R \cup (S^{\mathsf{tr}};R))y$ $(0 \leq m)$. Occurrences of aRc will be replaced by $aRbRc$. If $F_{n+1} \models z_i(S^{\mathsf{tr}};R)z_{i+1}$ but $F_n \models \neg(z_i(S^{\mathsf{tr}};R)z_{i+1})$, we consider several cases.

The case that $F_{n+1} \models b(S^{\mathsf{tr}};R)c$ but $F_n \models \neg(b(S^{\mathsf{tr}};R)c)$ is actually not possible and excluded by the other invariants. But still, in this hypothetical case, we could replace $b(S^{\mathsf{tr}};R)c$ by bRc.

In case $b = z_i(S^{\mathsf{tr}};R)z_{i+1} \neq c$ we transform it to bcz_{i+1} and note that now that either $F_n \models bRc(S^{\mathsf{tr}};R)z_{i+1}$ or $F_n \models bRcRz_{i+1}$.

In case $F_{n+1} \models b \neq z_i(S^{\mathsf{tr}};R)z_{i+1} \neq c$, we can replace $z_i z_{i+1}$ by $z_i c z_{i+1}$ Note that there are now two possibilities. Either $F_n \models z_i(S^{\mathsf{tr}};R)c(S^{\mathsf{tr}};R)z_{i+1}$ or $F_n \models z_i(S^{\mathsf{tr}};R)cRz_{i+1}$.

In case $F_{n+1} \models b \neq z_i(S^{\mathsf{tr}};R)z_{i+1} = c$ we see that necessarily $F_n \models z_i(S^{\mathsf{tr}};R)bRc$. Thus in this case, we can replace $z_i c$ by $z_i bc$. Note that indeed $F_n \models z_i(S^{\mathsf{tr}};R)bRc$ is necessary. For, suppose $F_{n+1} \models z_i(S^{\mathsf{tr}};R)c$ but $F_n \models \neg(z_i(S^{\mathsf{tr}};R)c)$. Thus, for any x such that $F_{n+1} \models z_i S^{\mathsf{tr}} xRc$, we have $F_n \models \neg(z_i S^{\mathsf{tr}} x) \vee negxRy$. If $F_n \models \neg xRc$, then $x = a$. It can not

[15] To be pedantically precise, we should write $\mathsf{S}^{\mathsf{tr,refl}}$, the transitive and reflexive closure of S.

be the case that also $F_n \models \neg(z_i S^{\mathsf{tr}} a)$. For, in this case we should have $F_n \models z_i S^{\mathsf{tr}} bRcS^{\mathsf{tr}} aRb$ which conflicts the conversely well-foundedness of $S^{\mathsf{tr}}; R^{\mathsf{tr}}$ on F_n. Thus $F_n \models z_i S^{\mathsf{tr}}$, whence $F_n \models z_i(S^{\mathsf{tr}}; R)bRc$.

In case $F_n \models \neg(z_i S^{\mathsf{tr}} x) \wedge xRc$, we get $F_n \models z_i S^{\mathsf{tr}} bRcS^{\mathsf{tr}} xRc$ or $F_n \models z_i S^{\mathsf{tr}} bRcRc$. This contradicts either the conversely well-foundedness of $S^{\mathsf{tr}}; R^{\mathsf{tr}}$ or the conversely well-foundedness of R.

In case $x = a$, that is, in case we consider $y \in \mathcal{M}_x^A$, we should also replace any occurrence of $bS_a c$ in the sequence $z_1, \ldots, z_l$ by bRc.

It is clear that the thus defined sequence witnesses $F_n \models y \in \mathcal{M}_x^A$.

Case (v): We have eliminated an imperfection concerning the $\mathbf{IL}\mathsf{M}$ frame-condition and $F_{n+1} := \langle W_n, R_n \cup \{\langle b, d\rangle\}, S_n, \nu_n\rangle$. We need to check all of the eight requirements. To see 1, the conversely well-foundedness of R, we reason as follows. Suppose for a contradiction that there is an infinite sequence such that $F_{n+1} \models x_1 R x_2 R \ldots$. We now get an infinite sequence $y_1, y_2, \ldots$ by replacing every occurrence of bRd in $x_1, x_2, \ldots$ by $bS_a cRd$ and leaving the rest unchanged. If there are infinitely many S_a-transitions in the sequence $y_1, y_2, \ldots$ (note that there are certainly infinitely many R-transitions in $y_1, y_2, \ldots$), 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.

Requirement 2 is easy as $S_{n+1} = S_n$. For 3 we only consider bRd. As we have $\nu(b) \subseteq_{\Box} \nu(c) \prec \nu(d)$, we also have $\nu(b) \prec \nu(d)$.

To see the validity of 4 and we reason as in the proof of Lemma 5.2 Actually, the proof goes completely the same. We only replace now occurrences of bRd in F_{n+1} by $bS_a cRd$ in F_n. A proof of 6 goes similarly to the proof of 6 in Case (i). Thus, a hypothetical infinite sequence such that $F_{n+1} \models x_1 S^{\mathsf{tr}} x_2 R^{\mathsf{tr}} x_3 S^{\mathsf{tr}} x_4 \ldots$ is transformed into an infinite sequence $y_1, y_2, \ldots$ by replacing bRd by $bS_a cRd$, yielding as always to a contradiction. Again, 7 is easy, as $S_{n+1} = S_n$.

To see $8'$ we reason as usual. We actually need not to make any replacements at all as $F_n \models b(S^{\mathsf{tr}}; R)d$. (Note that $S^{\mathsf{tr}}; (S^{\mathsf{tr}}; R) \subseteq S^{\mathsf{tr}}; R$.)

Thus, indeed F_n is a quasi-$\mathbf{IL}\mathsf{M}$ frame for any n. We will now 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 $F \models x_0 R x_1 R x_2 R \ldots$ defines an infinite sequence $F_0 \models x_0(S^{\mathsf{tr,refl}}; R^{\mathsf{tr}})x_1(S^{\mathsf{tr,refl}}; R^{\mathsf{tr}})x_2 \ldots$, 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 $(n.) :\ yS_x z \rightarrow \nu(y) \subseteq_{\Box} \nu(z)$. It is obvious that $(n.)$ holds on F. So, it remains to see that $(m.)$ holds.

But, as this has the status of a possible imperfection, we can just copy the proof from Lemma 5.2.

$\dashv$

Again do we note that the closure obtained in Lemma 6.5 is unique. Thus we can refer to the **IL**M-closure of a quasi-**IL**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 **IL**M-frames.[16]

Definition 6.6. An **IL**M-*quasi-frame* is defined as being that what you get from quasi-**IL**M-frames by dropping all the information about the labels. Thus, an **IL**M-quasi-frame is a triple $\langle W, R, S\rangle$ with W a non-empty set of worlds. R is a binary relation on W and S a set of binary relations on W indexed by elements of W. We have the following requirements.

1. $yS_xz \to xRy \ \& \ xRz$
2. R is conversely well-founded
3. $S^{\mathsf{tr}}; R^{\mathsf{tr}}$ is conversely well-founded

Corollary 6.7. *Any* **IL**M*-quasi-frame can be extended to an* **IL**M*-frame.*

Proof. One just has to copy the relevant parts of the proof of Lemma 6.5 $\dashv$

Corollary 6.8. *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-***IL**M*-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. The proof is as the proof of Corollary 5.3. We only need to remark on Case (v): If bS_acRd, 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 $A \notin \nu(b)$. $\dashv$

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

Lemma 6.9. *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 4.21 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 $D \wedge \Box E \wedge \Box\neg D \in \Delta''$. $\dashv$

[16]We note that not every **IL**W-frame can be extended to an **IL**M-frame. A counterexample is the smallest **IL**-frame containing $\langle\{a, b, c, d, e, f\}, \{\langle a, b\rangle, \langle b, c\rangle, \langle d, e\rangle, \langle e, f\rangle\}, \{\langle a, c, e\rangle, \langle d, f, b\rangle\}\rangle$.

6.2 Completeness

Theorem 6.10. **IL**M *is a complete logic.*

Proof. In Remark 4.22 four sufficient ingredients are mentioned for a logic to be complete. We give these ingredients in case **IL**X is **IL**M.

Frame Condition In the case of **IL**M the frame condition is easy and well known, as expressed in Lemma 6.2. Note that in the completeness proof we only use one part of the equivalence, that is, $F \models \forall x, y, u, v\ (yS_x uRv \to yRv) \Rightarrow F \models \mathbf{IL}\mathsf{M}$

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

   To see Requirement 8, $y \in \mathcal{M}_x^E \to \nu(x) \prec_E \nu(y)$, we reason as follows. Again, we first note that $\forall\, x, y{\in}W\ [G' \models y \in \mathcal{M}_x^C$ iff. $F \models y \in \mathcal{M}_x^C]$ holds for any C. We only need to consider the new element, that is, $b \in \mathcal{M}_x^E$. 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 $a' S^{\mathrm{tr}} a$. 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 6.8. 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 6.9 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 $yS^{\mathrm{tr}}a$, 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. $\dashv$

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 6.11.

(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) [17] $\mathbf{IL}\mathsf{M} \vdash A \rhd B \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \Diamond A \to \Diamond B$

[17]We thank Rosalie Iemhoff for pointing out this admissible rule to us.

(*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 $\mathsf{J4}$ 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 $r \notin W$)

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

It is easy to see that M' is an $\mathbf{IL}\mathsf{M}$-model. 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 $\mathbf{IL}\mathsf{M} \nvdash \Diamond A \to \Diamond B$.[18]

[18]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.

(v). The "$\Leftarrow$" direction is easy. For the other direction we reason as follows.[19]

We assume that $\nvdash A \rhd B$ and set out to prove $\nvdash \bigwedge \Diamond A_i \rightarrow A \rhd B$. 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 $\mathbf{IL}\mathsf{M} \vdash \Box\bot \rightarrow A$. 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 $\mathbf{IL}\mathsf{M} \nvdash \Box\bot \rightarrow A$. 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 6.11 we can combine (iii) and (iv) to obtain $\mathbf{IL}\mathsf{M} \vdash A \rightarrow B \vee \Diamond B \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \Diamond A \rightarrow \Diamond B$. 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 method. On the other hand, it is an indispensable ingredient in proving Theorem 7.4.

7 The essential Σ_1-sentences of essentially reflexive theories

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.

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

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 **IL**M. Our proof shall use their proof method.

We will perform our argument fully in **IL**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.

The result we present here seems extremely trivial. However, experience has taught us that matters concerning the complexity of arithmetic applications of modal formulae, tend to be very tricky. We should not forget, that we are talking a Σ_3-complete phenomenon ([Sha97]). However, in essentially reflexive theories, interpretability becomes a Π_2-matter. In a sense our result reflects the fact that the complexity of interpretability can not be reduced to Σ_1-phenomena.

7.1 Model construction

Throughout this section, 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 6.11 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 **GL**, we see that Theorem 7.2 also holds for **GL**.

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

Proof. It is well known that for essentially reflexive theories T we have that if σ is equivalent in T to a Σ_1!-sentence, then for all α and β

$$T \vdash \alpha \rhd \beta \to \alpha \wedge \sigma \rhd \beta \wedge \sigma.$$

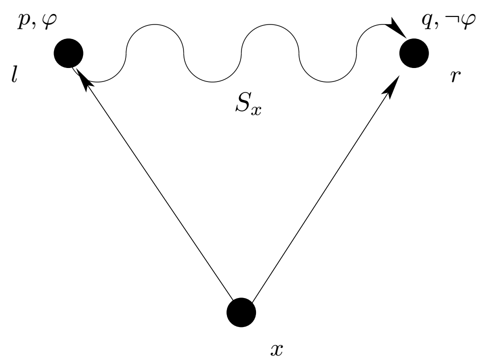

Figure 1: Counter model

If thus $\varphi \in \Sigma_1(T)$, we have that

$$\forall * \ T \vdash p^* \rhd q^* \to p^* \wedge \varphi^* \rhd q^* \wedge \varphi^*.$$

By Theorem 2.26 we see that $\mathbf{IL}\mathsf{M} \vdash p \rhd q \to p \wedge \varphi \rhd q \wedge \varphi$. $\dashv$

Theorem 7.4. *A formula φ is essentially Σ_1 iff. it is equivalent to some disjunction of $\Box$-formulas. That is, $\varphi \in \Sigma_1(T) \Leftrightarrow \mathbf{IL}\mathsf{M} \vdash \varphi \leftrightarrow \bigvee_{i\in I} \Box C_i$ for some $\{C_i\}_{i\in I}$.*

Proof. If $\mathbf{IL}\mathsf{M} \vdash \varphi \leftrightarrow \bigvee_{i\in I} \Box C_i$, then $\forall * \ T \vdash \varphi^* \leftrightarrow (\bigvee_{i\in I} \Box C_i)^*$ and clearly $(\bigvee_{i\in I} \Box C_i)^* \in \Sigma_1(T)$. We should thus concentrate on the other direction.

So we suppose that $\mathbf{IL}\mathsf{M} \nvdash \varphi \leftrightarrow \bigvee_{i\in I} \Box C_i$ for any finite set $\{C_i\}_{i\in I}$, and set out to prove that $\varphi \notin \Sigma_1(T)$. By Lemma 7.3 we are done if we can show that $\mathbf{IL}\mathsf{M} \nvdash p \rhd q \to p \wedge \varphi \rhd q \wedge \varphi$ for some p, $q \notin \varphi$. Or equivalently, by Theorem 6.10, we are done if we can expose a model $M \not\models p \rhd q \to p \wedge \varphi \rhd q \wedge \varphi$.

The idea is to build a model as in Figure 1, where p is true only in a world l (from left) and q only in a world r (right) and $l \Vdash \varphi$ and $r \Vdash \neg\varphi$. Moreover we shall have $xRlS_x r$. The $\mathbf{IL}\mathsf{M}$ frame-condition demands in this model $rRy \to lRy$. If our model M is to be an $\mathbf{IL}\mathsf{M}$-labeled $\mathbf{IL}\mathsf{M}$-model, we should thus certainly have $\nu(l) \subseteq_\Box \nu(r)$. If we moreover want a truth lemma to hold, we actually want $\varphi \in \nu(l) \subseteq_\Box \nu(r) \ni \neg\varphi$.

That we can find such labels for l and r is expressed by Lemma 7.9 (we call this the Σ-lemma). That having these labels is sufficient, follows from an application of the Main Lemma, Lemma 4.19, as we shall see. We shall not include the x in the labeled model. Rather shall we later place it under the rest of the model.

Our application of the Main Lemma runs as follows. We work under the assumption that φ is not equivalent to a disjunction of $\Box$-formulas. Let Δ_0 and Δ_1 be the MCS's as provided by Lemma 7.9. Thus, $\varphi \in \nu(l) \subseteq_\Box \nu(r) \ni \neg\varphi$. Let $\mathcal{D}$ be the smallest set of sentences that contains

φ and that is closed under taking subformulas and single negations. Let $\mathcal{I}$ be the following set of invariants.

$$\mathcal{I} := \begin{cases} xRy \rightarrow \exists\, A{\in}((\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}) \\ uS_x vRw \rightarrow uRw \\ \neg y(S \cup R)l \\ \forall\, x{\notin}\{l,r\}\ lR^{\mathsf{tr}}x \end{cases}$$

Note that the invariant $\forall\, x{\notin}\{l,r\}\ lR^{\mathsf{tr}}x$ certainly implies $rRy \rightarrow lRy$ if R is transitive. Clearly $\mathcal{I}^{\mathcal{U}} = \mathcal{I}$ and $G \models \mathcal{I} \Rightarrow G \models \mathbf{IL}\mathsf{M}$. We now define

$$F := \langle \{l,r\}, \varnothing, \varnothing, \{\langle l, \Delta_0\rangle, \langle r, \Delta_1\rangle\}\rangle. \qquad (\dagger)$$

As R and S are empty on F, F certainly is an $\mathbf{IL}\mathsf{M}$-frame on which our four invariants hold. We now want to apply the Main Lemma to extend F to some labeled adequate $\mathbf{IL}\mathsf{M}$-frame $\hat{F}$, on which a truth lemma holds with respect to $\mathcal{D}$. Thus, we should see that for any adequate labeled frame F', satisfying all the invariants, we can eliminate both deficiencies and problems, by extending F' in a way that conserves all the invariants.

We will copy these elimination processes in great part from the proof of Theorem 6.10. Thus, we consider an $\mathbf{IL}\mathsf{M}$-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 4.20 we can find a MCS Δ with $\nu(a) \prec_B \Delta \ni A, \Box\neg A$. We fix some $b \notin W$. By R^* we denote the reflexive closure of R, that is, $xR^*y \leftrightarrow x = y \vee xRy$. If $F' \models \neg(rR^*a)$, we define

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

   In case $F' \models rR^*a$, we define

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

   We should see that in both cases we get a quasi-$\mathbf{IL}\mathsf{M}$-frame on which all four invariants hold. In case $F' \models \neg(rR^*a)$ the argument is the same as in the proof of Theorem 6.10. We should only say a short word on the two new invariants. It is clear that $G' \models \neg y(S \cup R)l$, as $F' \models \neg y(S \cup R)l$. If we want to see that $G' \models \forall\, x{\notin}\{l,r\}\ lR^{\mathsf{tr}}x$, we only consider the new element b. But, $F' \models lRa$, whence $G' \models lRaRb$ and $G' \models lR^{\mathsf{tr}}b$.

   In case $F' \models rR^*a$, again it is easy to see that the invariants $\forall\, x{\notin}\{l,r\}\ lR^{\mathsf{tr}}x$ and $\neg y(S \cup R)l$ hold. We should now check the requirements that involve lRb.

   It is clear that R is conversely well-founded on G', as any occurrence of lRb or aRb in a sequence can only be at the very end of it.

   The only novelty is Requirement 3 of quasi-frames, $lRb \rightarrow \nu(l) \prec \nu(b)$. If rRa, then by the fact that F' satisfies all the invariants, lRa, whence $\nu(l) \prec \nu(a)$. But, by definition $\nu(a) \prec \nu(b)$ and thus $\nu(l) \prec \nu(b)$. If $r = a$, we see that $\nu(l) \subseteq_{\Box} \nu(r) \prec \nu(b) \rightarrow \nu(l) \prec \nu(b)$.

To see for example Requirement 4 of quasi-frames, $C \neq D \rightarrow \mathcal{G}_x^C \cap \mathcal{G}_x^D = \varnothing$, we reason as follows. We only have to consider the new element b. So, suppose $b \in \mathcal{G}_x^C \cap \mathcal{G}_x^D$. If $x = a$, then necessarily $C = B = D$. Thus, we may assume that $x \neq a$.

If $l = x$, either $\nu(\langle l, b\rangle) = C$, or $l \in \mathcal{G}_l^C$ or $a \in \mathcal{G}_l^C$ and similarly for D. However, the pair $\langle l, b\rangle$ is not labeled. Also $l \in \mathcal{G}_l^C$ is not possible. One can prove by induction on $\mathcal{G}_x^C$ that in (adequate labeled) quasi-**IL**M-frames $x \notin \mathcal{G}_x^C$.[20] Alternatively, we can see that $l \notin \mathcal{G}_l^C$ as our invariant $\neg y(R \cup S)l$ holds on G'. Consequently, $F' \models a \in \mathcal{G}_l^C \cap \mathcal{G}_l^D$ and $C = D$.

If $l \neq x \neq a$, it is clear that $a \in \mathcal{G}_x^C$ or $l \in \mathcal{G}_x^C$. Again $l \in \mathcal{G}_x^C$ is impossible, whence $F' \models a \in \mathcal{G}_l^C \cap \mathcal{G}_l^D$ and $C = D$. Requirement 8 admits a similar proof.

2. We now use Lemma 6.5 to extend G' to an adequate labeled **IL**M-frame G. We are to see that all the invariants hold on G. The only new invariants are $\neg y(S \cup R)l$ and $\forall\, x {\notin} \{l, r\}\ lR^{\mathrm{tr}}x$. As $lRaRb$, and as R is transitive on G, we see that lRb. Checking $\neg y(S \cup R)l$ should be done by taking it along the proof of Lemma 6.5, but this is completely trivial. Clearly $\langle a, \neg(A \rhd B)\rangle$ is not 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 2.8, or just by the fact that F' is a labeled **IL**M-frame, we have that $\nu(a) \prec_B \nu(b)$.

   By Lemma 6.9 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 $c \notin W$. If $F' \models \neg(rR^*a)$, we define
$$G' := \langle W, R \cup \{a, c\}, S \cup \{a, b, c\}, \nu \cup \{c, \Delta'\}\rangle.$$
   In case $F' \models rR^*a$, we define
$$G' := \langle W, R \cup \{\langle a, c\rangle, \langle l, c\rangle\}, S \cup \{\langle a, b, c\rangle\}, \nu \cup \{\langle c, \Delta'\rangle\}\rangle.$$
   In the first case we see by an argument as in the proof of Theorem 6.10 that G' is a quasi-**IL**M-frame. Again, it is easy to see that G' satisfies the two invariants involving l.

   In the second case we should reason the same as earlier in this proof, when we eliminated a problem, to see that G' is a quasi-**IL**M-frame satisfying all the invariants.

2. We close G' off under the frame conditions with Lemma 6.5. It is a routine inspection to see that all the invariants hold on G. Clearly, $\langle a, b, C \rhd D\rangle$ is not a deficiency on G.

Thus, we now have shown that any problem or deficiency in any adequate labeled frame F' satisfying all the invariants can be eliminated

[20] This does not hold in all **IL**X-frames.

by extending F'. Moreover, we can do that in such a way that all the invariants hold on the extension. By the Main Lemma, we get that consequently any labeled frame, satisfying all the invariants can be extended to an **IL**M-frame on which a truth lemma with respect to $\mathcal{D}$ holds.

Consequently, the F that we defined in (†) can be extended to some $\hat{F}$ on which a truth lemma holds w.r.t. $\mathcal{D}$. Let $\overline{F} = \langle \overline{W}, \overline{R}, \overline{S}, \overline{\Vdash} \rangle$ be the model induced by $\hat{F}$. Certainly we have that

$$\overline{F}, l \Vdash \varphi \quad \& \quad \overline{F}, r \Vdash \neg\varphi.$$

We transform $\overline{F}$ into our required **IL**M-model M by "gluing" a root w_0 to it ($w_0 \notin \overline{F}$):

$$\begin{aligned} M := \quad & \langle W \cup \{w_0\}, R \cup \{\langle w_0, w\rangle \mid w \in \overline{W}\}, \\ & S \cup \{\langle w_0, x, y\rangle \mid \langle x, y\rangle \in \overline{R} \text{ or } x{=}y \in \overline{W}\} \cup \{\langle w_0, l, r\rangle\}, \Vdash_M\rangle. \end{aligned}$$

Let p and q be propositional variables that do not yet occur in the range of $\overline{\Vdash}$. We define $\Vdash_M$ as follows.

$$M, x \Vdash_M s \quad \text{iff.} \quad \begin{cases} x{\neq}w_0 \ \& \ s \notin \{p, q\} \ \& \ \overline{F}, x \overline{\Vdash} s \text{ or} \\ x{=}l \ \& \ s{=}p \text{ or} \\ x{=}r \ \& \ s{=}q \end{cases}$$

It is easy to check that M is an **IL**M-model. Also, using Lemma 2.12, it is clear that $M, w_0 \Vdash \neg(p \rhd q \to p \wedge \varphi \rhd q \wedge \varphi)$. ⊣

The proof we have just presented gives us some additional information. For a set $\mathcal{D}$, we can define the following relation on MCS's.

$$\Gamma \prec^{\mathcal{D}} \Delta \ \text{ iff. } \ \Box A \in \Gamma \cap \mathcal{D} \to A, \Box A \in \Delta$$

Corollary 7.5. *Let Δ_0, Δ_1 and $\mathcal{D}$ be as in the proof above of Theorem 7.4. There is a* MCS Γ *such that* $\Gamma \prec^{\mathcal{D}} \Delta_0, \Delta_1$.

Proof. By inspection of the proof of Theorem 7.4 and using Remark 4.5. The label that m_0 gets in $\overline{M}$ will be the required Γ. ⊣

It is not hard to get generalizations of Corollary 7.5, like Lemma 7.6, by slightly changing our Main Lemma. But these sort of statements are often also easy to obtain directly.

Lemma 7.6. *Let Δ_0 and Δ_1 be maximal* **IL**M*-consistent sets. There is a maximal* **IL**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 6.11 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. ⊣

Actually, we can take Lemma 7.6 as a starting point for an alternative proof of Theorem 7.4. The idea of the proof does not change a lot and is still captured in Picture 1. Thus, given some φ that is not equivalent to any disjunction of $\Box$-formulas, we use Δ_0, Δ_1 and Γ as given by Lemma's 7.9 and 7.6, and define

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

We then take the obvious $\mathcal{D}$ and I and apply the Main Lemma. The benefit of this alternative approach is that we can use that part of the proof of Theorem 6.10 that concerns the elimination of problems and deficiencies.

Remark 7.7. There is however one subtlety in this alternative approach. One of the invariants in the Main Lemma required that we could find some specific $\Box$-formula in $\nu(y)\backslash\nu(x)$ whenever xRy. How can we possibly guarantee that there is some $\Box\chi \in \Delta_0\backslash\Gamma$? It might very well be the case that $\chi \in \Delta_0 \Rightarrow \Diamond\chi \in \Delta_0$ for all χ. We see various ways out.

- We could add a new propositional constant symbol c to the language and extend Δ_0 to $\Delta_0' := \Delta_0 \cup \{c, \Box\neg c\}$. This Δ_0' is again consistent. For, if $\Delta_0, c, \Box\neg c \vdash \bot$, then for some $\delta \in \Delta_0$ we get $\delta \vdash \Box\neg c \to \neg c$, whence $\vdash\to \delta(\Box\neg c \to \neg c)$. As $\mathbf{IL}\mathsf{M}$ is closed under substitution, we see that $\vdash \delta \to (\Box\neg\delta \to \neg\delta)$, whence $\vdash \Box\neg\delta \to \neg\delta$ and by Löb $\vdash \neg\delta$. This contradicts the consistency of Δ_0.

  Similarly, we can extend Δ_1 to Δ_1'. But, now it is still a tour de force to get maximal consistent extensions $\widetilde{\Delta_0} \subseteq_\Box \widetilde{\Delta_1}$ in the enriched language.

- We could try to strengthen Lemma 7.9 to get a more informative lemma. This more informative lemma should give us $\varphi, \Box\chi \in \Delta_0 \subseteq_\Box \Delta_1 \ni \neg\varphi$. Really, some work has to be done here, as all the obvious attempts seem to fail. For example, if $\varphi = A \rhd B$, we can never get $\varphi, \Box\neg\varphi \in \Delta_0$. (For then, $\neg\varphi, \Box\bot \in \Delta_1$, which is impossible.) In this case, we can however get $\varphi, \Box\varphi \in \Delta_0 \subseteq_\Box \Delta_1 \ni \neg\varphi$. For $\varphi = \neg(A \rhd B)$ we can not get $\varphi, \Box\varphi \in \Delta_0$. But in this case we could work with $\varphi, \Box\neg\varphi \in \Delta_0 \subseteq_\Box \Delta_1 \ni \neg\varphi$. It seems tempting to conjecture that we can either use $\varphi, \Box\varphi \in \Delta_0$ or $\varphi, \Box\neg\varphi \in \Delta_0$. This however is not the case as we shall show in Section 9.4. A good analysis for which formulas we can do this trick seems equally hard as the problem we originally started out with.

  However, in a generalization of Lemma 7.9 we need not necessarily use either $\chi = \varphi$ or $\chi = \neg\varphi$ to get $\varphi, \Box\chi \in \Delta_0 \subseteq_\Box \Delta_1 \ni \neg\varphi$.

- We could generalize the Main Lemma. The proof of the Main Lemma still goes through if we have $xRy \to \exists\, A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box\neg D \mid D \text{ a subformula of some } B \in \mathcal{D}\}$ for all x, y where $x \neq m_0$. We would still get a bound on our chains.

The alternative proof that we give of Theorem 7.4 is based on a generalization of the Main Lemma.

Lemma 7.8 (Generalized 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 and let* G *be an adequate* $\mathbf{IL}\mathsf{X}$*-labeled frame. Moreover, let* $\mathsf{G} \subseteq G$ *be a finite subset of the universe of* G*. Let* $\mathcal{I}$ *be a set of so-called* invariants *of labeled frames so that we have the following properties. In the following,* F *is an arbitrary extension of* G*.*

- $G \models \mathcal{I}$
- $F \models \mathcal{I}^{\mathcal{U}} \Rightarrow F \models \mathcal{C}$*, 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:* $xRy \wedge x \notin \mathsf{G} \rightarrow \exists\, A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box\neg D \mid D \text{ a subformula of some } B \in \mathcal{D}\}$.
- *If* F *satisfies all the invariants, then 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 this case, we can extend G *to an adequate labeled* $\mathbf{IL}\mathsf{X}$*-frame* $\hat{G}$ *on which a truth lemma holds with respect to* $\mathcal{D}$*.*

It is obvious how to modify the proof of the Main Lemma, so to obtain a proof of the Generalized Main Lemma.

Second Proof of Theorem 7.4. Let φ be a formula that is not equivalent to a disjunction of $\Box$-formulas. According to Lemma 7.9 we can find MCS's Δ_0 and Δ_1 with $\varphi \in \Delta_0 \subseteq_{\Box} \Delta_1 \ni \neg\varphi$. By Lemma 7.6 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 the Generalized Main Lemma to this frame G. The finite set G will be just $\{m_0\}$. 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.

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

In the proof of Theorem 6.10 we have seen that we can eliminate both problems and deficiencies while conserving the invariants. The Generalized 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 $m_0 \Vdash \neg(p \rhd q \rightarrow p \wedge \varphi \rhd q \wedge \varphi)$, whence by Lemma 7.3 we get $\varphi \notin \Sigma_1(T)$.

⊣

7.2 The Σ-lemma

We can say that the proof of Theorem 7.4 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 7.9. *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 $Y \in \Box_{\text{con}}$ a finite set $Y^{\text{fin}} \subseteq \Box_{\vee} - Y$ such that

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

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

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

Now we will show that:

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

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 $\Box^{\mathrm{fin}}_{\mathrm{con}} \subseteq \Box_{\mathrm{con}}$ such that

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

By (3) we also have

$$\vdash \varphi \to \bigwedge_{Y \in \Box^{\mathrm{fin}}_{\mathrm{con}}} \bigvee_{\sigma \in Y^{\mathrm{fin}}} \sigma. \tag{6}$$

Combining (5) with (6) we get

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

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 **IL**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.

By Lemma 9.6 and by some elementary semantical argument, we see 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.

8 Essentially Σ_1-sentences for reasonable arithmetical theories

The content of this section is mainly negative. We show that we can not directly generalize our proof method for showing essentially Σ_1-ness in essentially reflexive theories, to any reasonable arithmetical theory. There is also some positive information contained though in this section. We give a necessary condition for a formula to be Σ_1 and we conjecture a characterization of the Σ_1-formulas of $\mathbf{IL}(\mathrm{All})$. Also do we embark on the relation between interpretability and provability in finitely axiomatized theories.

8.1 A necessary condition for Σ_1-ness

We want a necessary condition for a sentence to be Σ_1 to perform a reasoning as in Theorem 7.4. Provable Σ_1-completeness, that is $\sigma \to \Box\sigma$ for Σ_1-formulas σ, is such a condition. The condition we look for should however not hold for all so-called self provers as studied by Kent ([Ken73]) and Guaspari ([Gua83]). Notably the formula $p \wedge \Box p$ should not satisfy the necessary condition. The following lemma provides such a condition.

Lemma 8.1. *Let T be a reasonable arithmetical theory. If $\sigma \in \Sigma_1(T)$, then $\forall\alpha\forall\beta\ T \vdash \alpha \rhd \beta \to \Diamond\alpha \wedge \sigma \rhd \beta \wedge \sigma$.*

Proof. The proof of this lemma is basically the soundness proof of principle M_0. We may assume that $\sigma = \exists x\ \varphi(x)$, with $\varphi \in \Delta_0$.

We reason in T and assume that for some interpretation $j : \alpha \rhd \beta$. This j comes with a special $T + \alpha$ cut J and a definable isomorphism F. The F is an isomorphism between J, living in some model M of $T + \alpha$, and an initial segment of the model M' of $T + \beta$ that is internally defined in M by our interpretation j.

We want to see that $\Diamond\alpha \wedge \sigma \rhd \beta \wedge \sigma$. To this extend we assume $\Diamond\alpha \wedge \sigma$ which is the same as $\Diamond\alpha \wedge \exists x\ \varphi(x)$. By Lemma 8.2 this implies $\Diamond\alpha \wedge \exists x\ \Box(x \in J \wedge \varphi(x))$. Thus, we also have $\Diamond\alpha \wedge \Box\exists x\ (x \in J \wedge \varphi(x))$ and $\Diamond(\alpha \wedge \exists x\ (x \in J \wedge \varphi(x)))$.

Consequently, $\Diamond\alpha \wedge \exists x\ \varphi(x) \rhd \Diamond(\alpha \wedge \exists x\ (x \in J \wedge \varphi(x)))$ and by $\mathsf{J5}$ and $\mathsf{J2}$, also $\Diamond\alpha \wedge \exists x\ \varphi(x) \rhd \alpha \wedge \exists x\ (x \in J \wedge \varphi(x))$. By the choice of J and F and by the fact that φ is a Δ_0-formula, we now see $\alpha \wedge \exists x\ (x \in J \wedge \varphi(x)) \rhd \beta \wedge \exists x\ \varphi(x)$. Thus we have proved $\Diamond\alpha \wedge \sigma \rhd \beta \wedge \sigma$. $\dashv$

Lemma 8.2. *Let T be a reasonable arithmetical theory. Let J be a T-cut and let $\varphi(x) \in \Delta_0$. We have that $T \vdash \exists x\ \varphi(x) \to \exists x\ \mathsf{Prov}_T(x \in J \wedge \varphi(x))$.*

Proof. It is well known that $T \vdash \forall x\ \mathsf{Prov}_T(x \in J)$. For theories where the exponentiation function is total, we get a very easy proof. For weaker theories one has to switch to efficient coding techniques. See for example [JV00]. The principle is sometimes called the "outside big, inside small" principle. We write $\mathcal{D}$ for the part of the standard proof such that

$$\begin{array}{c}\mathcal{D}\\ \forall x\ \mathsf{Prov}_T(x \in J).\end{array}$$

To get our result, we reason in T as follows.

$$\dfrac{\dfrac{\dfrac{[\exists x\ \varphi(x)]_1 \quad \dfrac{\dfrac{[\varphi(c)]_2}{\mathsf{Bew}_T(\varphi(c))}}{\exists x\ \mathsf{Bew}_T(\varphi(x))}}{\exists x\ \mathsf{Bew}_T(\varphi(x))}\ \exists E, 2 \qquad \begin{matrix}\mathcal{D}\\ \forall x\ \mathsf{Prov}_T(x \in J)\end{matrix}}{\exists x\ \mathsf{Bew}_T(x \in J \wedge \varphi(x))}}{\exists x\ \varphi(x) \to \exists x\ \mathsf{Bew}_T(x \in J \wedge \varphi(x))}\ \to I, 1$$

⊣

8.2 Finitely axiomatized theories

In Section 7 we proved a characterization of essentially Σ_1-sentences for essentially reflexive theories. The proof used modal techniques only. We will now see that it is unlikely to find such a proof for finitely axiomatized theories.

We first note, that we can never use Lemma 8.1 to substitute Lemma 7.3 in the proof of Theorem 7.4. For, if we are to use the same proof strategy to show that $\neg\Sigma(\varphi)$, we are to come up with a labeled model M and some $a, b, c, d \in M$ such that the following hold.

- $aRbRcS_a d$
- $\varphi \in \nu(b) \subseteq_\Box \nu(d) \ni \neg\varphi$
- A truth lemma holds on M with respect to the smallest set containing φ and being closed under taking subformulas and single negations.

If this is possible, we can define, for some fresh p and q, that p is only true at c, and that q is only true at d. In such a case, $M, a \Vdash p \rhd q \wedge \neg(\Diamond p \wedge \varphi \rhd q \wedge \varphi)$, whence by Lemma 8.1 we see that $\neg\Sigma(\varphi)$. This approach however, does succeed if and only if $\nvdash \varphi \to \Box\bot$.

If $\vdash \varphi \to \Box\bot$, then, by our assumption that $M, b \Vdash \varphi$, also $M, b \Vdash \Box\bot$. As bRc, clearly $M, b \Vdash \Diamond\top$. This is impossible.

If $\nvdash \varphi \to \Box\bot$, then $\{\varphi, \Diamond\top\}$ is consistent. We can find $\varphi, \Diamond\top \in \Delta_0 \subseteq_\Box \Delta_1 \ni \neg\varphi$ by a lemma similar to Lemma 7.9. We now take any[21] Γ with $\Delta_0 \prec \Gamma$ and any $\Gamma_0 \prec \Delta_0$ and define

$$\begin{aligned} F := \quad & \langle \{a, b, c, d\}, \{\langle a, b\rangle, \langle a, c\rangle, \langle a, d\rangle, \langle b, c\rangle\}, \\ & \{\langle a, c, d\rangle\}, \{\langle a, \Gamma_0\rangle, \langle b, \Delta_0\rangle, \langle c, \Gamma\rangle, \langle d, \Delta_0\rangle\}\rangle. \end{aligned}$$

An application of the Generalized Main Lemma, Lemma 7.8, now yields the required model M.

The restriction that $\nvdash \varphi \to \Box\bot$ is however a serious restriction. There do exist φ with $\neg\Sigma(\varphi)$ and $\vdash \varphi \to \Box\bot$. The most prominent example is probably $\varphi = p \wedge \Box\bot$. Thus, probably, the pure modal formalism of $\mathbf{IL}\mathsf{P}$ is not refined enough to yield a characterization of essentially Σ_1-formulas

[21] Note that by Lemmas 4.20 and 7.6, we can indeed find such Γ and Γ_0.

of finitely axiomatized theories. A proof of Conjecture 8.3 should either use a richer modal signature, or a direct embedding of formulas φ that are not essentially Σ_1 into arithmetic. The latter approach has been applied in [Bek93].

Conjecture 8.3. *For finitely axiomatized theories we have that $\Sigma(\varphi)$ iff.* $\mathbf{IL}\mathsf{P} \vdash \varphi \leftrightarrow \bigvee\!\!\!\bigvee_i \Box A_i \vee (\bigvee\!\!\!\bigvee_j \bigwedge\!\!\!\bigwedge_k B_{jk} \rhd C_{jk})$ *for some suitable formulas and indices.*

It is clear that indeed all formulas of the form $\bigvee\!\!\!\bigvee_i \Box A_i \vee (\bigvee\!\!\!\bigvee_j \bigwedge\!\!\!\bigwedge_k B_{jk} \rhd C_{jk})$ are provably Σ_1 in any finitely axiomatizable theory. The other direction needs a special treatment.

In some respect, the logic $\mathbf{IL}\mathsf{P}$ seems to be an easier logic than $\mathbf{IL}\mathsf{M}$. First of all, the $\rhd$ in $\mathbf{IL}\mathsf{P}$ is used to describe a Σ_1-phenomenon. In $\mathbf{IL}\mathsf{M}$ interpretability describes a Π_2-matter. Second, we have a reduction from $\mathbf{IL}\mathsf{P}$ to $\mathbf{IL}$ in the sense of Lemma 8.5. Third, we see by Theorem 8.6 that in a certain sense, $\mathbf{IL}\mathsf{P}$ essentially reduces to provability logics.

Definition 8.4. We define a translation $\sharp$ of formulas to formulas, as follows. We define $p^\sharp = p$ for propositional variables. Moreover, $\sharp$ commutes with the boolean connectives and the $\Box$-modality. The only non-trivial action is on the $\rhd$-modality. In this case we define $(A \rhd B)^\sharp = A^\sharp \rhd B^\sharp \wedge \Box(A^\sharp \rhd B^\sharp)$.

Lemma 8.5 (Hájek). $\mathbf{IL}\mathsf{P} \vdash \varphi \Leftrightarrow \mathbf{IL} \vdash \varphi^\sharp$

Theorem 8.6. *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 $T \vdash ((\alpha \rhd_T \beta) \vee \Box_T \bot) \leftrightarrow \alpha \rhd_T \beta$. $\dashv$

In some sense Theorem 8.6 suggests that $\mathbf{IL}\mathsf{P}$ can be related to provability in a strong sense. The following definition indicates one of the lines amongst which one could work.

Definition 8.7. The logic $\mathbf{BIL}\mathsf{P}$ is given by its language, its axioms and its rules.

- The language of $\mathbf{BIL}\mathsf{P}$ is an extension of the language of $\mathbf{IL}\mathsf{P}$ such that for any $A \rhd B$ in the language of $\mathbf{BIL}\mathsf{P}$, there is a constant $P_{A \rhd B}$ in the language of $\mathbf{BIL}\mathsf{P}$.
- The axioms are just all the axioms of $\mathbf{IL}$, plus $A \rhd B \leftrightarrow \Box P_{A \rhd B}$, for any A, B in the language of $\mathbf{BIL}\mathsf{P}$.

- The rules are Modus Ponens and Necessitation

Now, if $\natural$ is a translation that replaces in φ, all occurrences of $A \rhd B$ by $\Box P_{A\rhd B}$ (inside out), and if $\mathbf{IL}\mathsf{P} \vdash \varphi$ iff. $\mathbf{BIL}\mathsf{P} \vdash \varphi^{\natural}$, we can relate Conjecture 8.3 to a conjecture in $\mathbf{BIL}\mathsf{P}$, and we are back with $\Box$-formulas.

8.3 Essentially Σ_1 in $\mathbf{IL}(\mathrm{All})$

If we compare Conjecture 8.3 to Theorem 7.4, we see that there can not be something as a characterization for essentially Σ_1-sentences that holds for any reasonable arithmetical theory. Both essentially reflexive theories, and finitely axiomatizable theories (of some reasonable strength) are reasonable arithmetical theories. Theorem 7.4 does however suggest a conjecture for the essentially Σ_1-sentences of $\mathbf{IL}(\mathrm{All})$.

Definition 8.8. A modal formula φ is an essentially Σ_1-sentence of $\mathbf{IL}(\mathrm{All})$ if $\forall T \forall *\ T \vdash \Sigma(\varphi^*)$.

Conjecture 8.9. *The essentially Σ_1-sentences of $\mathbf{IL}(\mathrm{All})$ are precisely those equivalent in $\mathbf{IL}(\mathrm{All})$ to a disjunctions of $\Box$-formulas.*

It is clear that a disjunction of $\Box$-formulas is indeed Σ_1 in any reasonable arithmetical theory T. The reasoning in $\mathbf{IL}(\mathrm{All})$ showing that a formula φ is equivalent to such a disjunction is by definition of $\mathbf{IL}(\mathrm{All})$ available in T. It seems natural to expect that the Σ_1-sentences of $\mathbf{IL}(\mathrm{All})$ can somehow be related to Theorem 7.4.

9 Self provers and formulas that generate trivial self provers

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*.

9.1 Formulas that generate trivial self provers

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 9.1. 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 $\mathbf{GL}$. In the rest of this subsection, $\vdash$ will stand for derivability in $\mathbf{GL}$. We shall often write Σ instead of Σ_1.

Theorem 9.2. *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 $\vdash \varphi\wedge\Box\varphi \leftrightarrow \bigvee\!\!\!\bigvee_m \Box C_m$. 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. By remark 9.3 we see that 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

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

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 2. 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 $L, w\Vdash\neg\Sigma(\varphi\wedge\Box\varphi)$. By the completeness of $\Sigma-$**IL**M and $\Sigma-$**GL** (see [Gor03]) we conclude that indeed $\neg\Sigma(\varphi\wedge\Box\varphi)$. $\dashv$

Remark 9.3. It is not hard to see that for any φ, we can find φ_l, A_l and C_m such that 1-3 hold. Just by propositional logic we can write φ in a way such that 1 and 3 hold. Now we can just run through φ and any time, if for a disjunct $\varphi_l\wedge\Box A_l$ we have that $\vdash \Box A_l\to\varphi$, we can

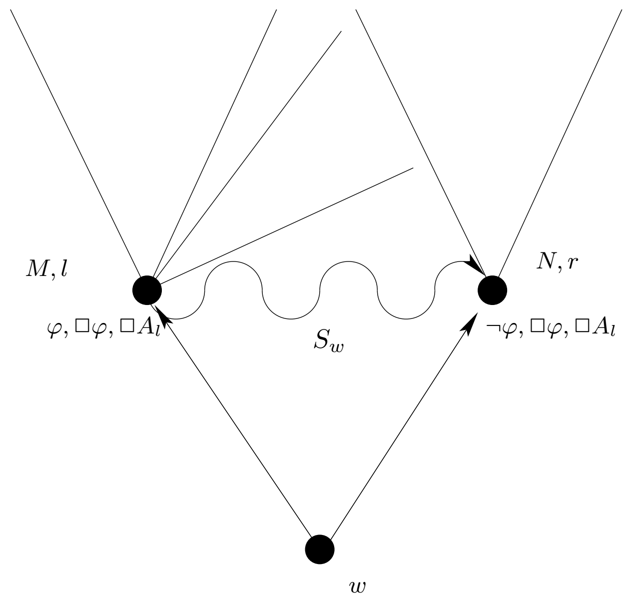

Figure 2: T.s.g.'s

replace that disjunct by $\Box A_l$ and obtain an equivalent formula. Clearly, this process comes to an end, and the final result satisfies 1-3. If we abbreviate $\bigvee\!\!\!\bigvee_{l' \neq l}(\varphi_{l'} \wedge \Box A_{l'}) \vee \bigvee\!\!\!\bigvee_m \Box C_m$ by Q_l, we see that the following are equivalent.

- $\vdash \Box A_l \rightarrow \varphi$
- $\vdash \Box A_l \rightarrow \varphi_l \vee Q_l$
- $\vdash \Box A_l \vee Q_l \leftrightarrow (\varphi_l \wedge \Box A_l) \vee Q_l$
- $\vdash \neg Q_l \wedge \Box A_l \rightarrow \varphi_l$

As we have seen, there do indeed exist t.s.g.'s that are not Σ_1. Actually, we can come up with quite some examples that do not look alike.

- $(\Box q \rightarrow \Box\Box q) \wedge (\Box\Box p \rightarrow \Box p)$
- $(\Box(q \wedge \Box r) \rightarrow \Box r) \wedge (\Box\Box q \rightarrow \Box q)$
- $(\Box p \rightarrow \Box q) \wedge (\Box q \rightarrow \Box p) \wedge (\Box\Box p \rightarrow \Box p)$
- $(\Box\Box q \rightarrow \Box q) \wedge (\Box\varphi \rightarrow \Box q \vee \Box\chi)$

We see that Löb's axiom plays a crucial role in showing that $\Sigma(\varphi \wedge \Box\varphi)$. Therefore, it seems natural to work with a sort of conjunctive normal form for φ.

We will now pronounce a conjecture of a characterization of $\Sigma(\varphi \wedge \Box\varphi)$ when φ is given in "conjunctive normal form".

Conjecture 9.4. *We have in* **GL** *that* $\Sigma(\varphi \wedge \Box\varphi)$ *iff. the following condition is satisfied.*

For all formulae φ_i, B_{ij} and C_k satisfying i, ii and iii, we have that $\vdash \varphi \wedge \Box\varphi \leftrightarrow (\bigvee_k \Box C_k) \wedge \bigwedge_i \bigvee_j \Box B_{ij}$*. Here i-iii are the following conditions.*

i. $\vdash \varphi \leftrightarrow \bigwedge_i(\varphi_i \vee \bigvee_j \Box B_{ij}) \wedge (\bigvee_k \Box C_k)$

ii. for all i we have $\nvdash Q_i \to \varphi_i \vee \bigvee_j \Box B_{ij}$
where $Q_i := \bigwedge_{i' \neq i}(\varphi_{i'} \vee \bigvee_j \Box B_{i'j}) \vee \bigvee_k \Box C_k$

iii. φ_i is a non-empty disjunction of literals and diamond-formulas

If we combine Conjecture 9.4 with Conjecture 9.14, we get that we can take the φ_i in Conjecture 9.4 to consist of just a single $\Diamond$-formula. Thus, the normal form of φ would look like $\bigwedge_i(\Box E_i \to \bigvee_j \Box F_{ij})$. Now, how can $\varphi \wedge \Box\varphi$ be possibly Σ_1? The only sensible reason can be that either $\Box E_i$ or $\bigvee_j \Box F_{ij}$ is provable for any i from $\varphi \wedge \Box\varphi$.[22] In either case $\bigvee_j \Box B_{ij}$ is provable for any i.

We conjecture that Theorem 9.2 and Conjecture 9.4 also hold for **IL**M. A proof can probably be given along the same lines. The only difference will be in the model construction in the proof of Theorem 9.2. In the case of **GL**, we could just merge two rooted models and glue a new common root to them. In the case of **IL**M, probably an application of the Main Lemma is required at that place.

9.2 Decidability of the problems "being Σ_1" and "being a t.s.g." in GL

In this subsection, we restrict our attention again to the logic **GL**. A great part of the discussion can be generalized to **IL**M though.

We know that a formula φ is essentially Σ_1 in **GL** if and only if it is equivalent to a disjunction of $\Box$-formulas. Formulated as such, it is prima facie not clear whether the notion of Σ_1-ness is decidable, as we have an unrestricted existential quantifier in the definition. It is very likely that this quantifier can be restricted to a finite class that is generated by the subformulas of φ. Below we shall give a model theoretic characterization of "being equivalent to a disjunction of $\Box$-formulas".

From this characterization, the decidability will follow. The decidability is actually a well-known fact and follows for example from work from Ignatiev ([Ign93]), Japaridze ([Jap94]) and Goris [Gor03].

Definition 9.5. Let M, r be a rooted tree for **GL**, with root r and $M := \langle W, R, \Vdash\rangle$. We say that a subtree $N, r := \langle W', R', \Vdash'\rangle$ of M, r is a *prune* of M, r if the following holds.

1. $r \in W' \subseteq W$
2. If $x \in W \setminus W'$ and $M \models xRy$ then $y \in W \setminus W'$
3. $x \in W \setminus W' \to \exists y{\in}W\setminus W'\ (M \models rRy \wedge \forall z\ \neg(rRzRy))$
4. $R' = R{\restriction}W'$
5. $x \in W' \setminus \{r\} \to (x \Vdash p$ iff. $x \Vdash' p)$ for all propositional variables p.

[22] It seems highly implausible that the implication can be provable without the antecedent or the succedent being provable.

We say that a formula φ is preserved under prunes if φ holds at the root of any prune of M, r whenever it holds in M, r. That is, $M, r \Vdash \varphi \Rightarrow N, r \Vdash \varphi$ for any prune N, r of M, r.

Lemma 9.6. *A modal formula φ is equivalent (in* **GL***) to a disjunction of $\Box$-formulas iff. φ is preserved under prunes.*

Proof. The $\Rightarrow$ direction is rather easy. It follows from Lemma 2.12 and the observation that generated submodels are preserved under taking prunes. That is, the generated submodel of x in N, r is the same as the generated submodel of x in M, r, whenever N, r is a prune of M, r and $x \neq r$.

For the $\Leftarrow$ we reason as follows. Suppose that φ is not equivalent to the disjunction of $\Box$-formulas. In other words (symbols), $\neg\Sigma(\varphi)$. By Theorem 10.5 we can find a Σ-**GL** model $L, w \Vdash \neg\Sigma(\varphi)$. In other words, we can find an **IL**M-model $L := \langle W, R, S, \Vdash\rangle$ as depicted in Figure 1 with $L, l \Vdash \varphi$ and $L, r \Vdash \neg\varphi$ and $L \models rRx \rightarrow lRx$. We now define $P, l := \langle W', R', \Vdash'\rangle$ with

$$\begin{array}{lcl} W' & := & \{x \in W \mid x = l \vee lRx\} \\ R' & := & R{\restriction}W' \\ \Vdash' & := & \Vdash{\restriction}W' \end{array}$$

and $Q, r := \langle W'', R'', \Vdash''\rangle$ with

$$\begin{array}{lcl} W'' & := & \{x \in W \mid x = r \vee rRx\} \\ R'' & := & R{\restriction}W'' \\ \Vdash'' & := & \Vdash{\restriction}W''. \end{array}$$

As L is an **IL**M-model, and thus $L \models lRx \rightarrow rRx$, we see that Q, r is a prune of P, l. $\dashv$

Corollary 9.7. *$\varphi \in \Sigma(T)$ is decidable in* **GL***.*

Proof. Enumerate for every $\bigvee\!\!\!\!\bigvee_i \Box C_i$ all possible proofs of $\varphi \leftrightarrow \bigvee\!\!\!\!\bigvee_i \Box C_i$. "Intertwine" this enumeration with checks of Lemma 9.6 for φ on all finite tree models with their corresponding prunes. $\dashv$

It is not clear what the complexity of this algorithm is. But probably, the restriction of the possible $\Box C_i$ to prove $\bigvee\!\!\!\!\bigvee_i \Box C_i \leftrightarrow \varphi$ is a lot more efficient.

In Subsection 9.1 we proved a characterization in **GL** for φ to be a t.s.g. We now address the question whether this is an informative characterization. As we have just seen, it is decidable whether $\Sigma(\varphi \wedge \Box\varphi)$. The characterizations we gave do not seem to give nicer complexity bounds on the problem of $\Sigma(\varphi \wedge \Box\varphi)$, as they are formulated using a universal quantifier. We do have however the following easy corollary to Theorem 9.2.

Corollary 9.8. *For all φ_l, A_l and C_m satisfying 1, 2 and 3 from Theorem 9.2 we have $\vdash \varphi \wedge \Box\varphi \leftrightarrow \bigvee\!\!\!\!\bigvee_m \Box C_m$*
iff.
There exist φ_l, A_l and C_m satisfying 1, 2 and 3 from Theorem 9.2 and $\vdash \varphi \wedge \Box\varphi \leftrightarrow \bigvee\!\!\!\!\bigvee_m \Box C_m$.

Proof. The "For all" implies the "There exist" direction is obvious in the light of Remark 9.3. For the other direction, suppose we have A_l, φ_l and C_m satisfying the requirements. Then clearly $\Sigma(\varphi \wedge \Box\varphi)$, whence by Theorem 9.2 we get the universal statement. $\dashv$

Corollary 9.9. *It is decidable if φ is a t.s.g. or not.*

Proof. Transform φ into "disjunctive normal form", that is, satisfying conditions 1, 2 and 3 from Theorem 9.2, and check whether $\vdash \varphi \leftrightarrow \bigvee\!\!\!\bigvee_m \Box C_m$. The correctness of this procedure is guaranteed by Remark 9.3, Theorem 9.2 and Corollary 9.8. $\dashv$

We now see that this procedure indeed looks a lot easier than the procedure we sketched for a direct check if $\Sigma(\varphi \wedge \Box\varphi)$. We actually see that the problem of φ being a t.s.g. when φ is given so that it satisfies 1-3 from Theorem 9.2 is in PSPACE. It can just be reduced to theoremhood in **GL**, which is PSPACE-complete. It is not hard to see that the problem is actually PSPACE-complete. Thus, indeed our characterizations are informative characterizations. Improvements on Conjecture 9.4 could probably be made by requiring additional direct recognizability of reflection principles. A reflection principle is a formula of the form $\Box\chi \to \chi$.

9.3 Formulas that generate modalized self provers in GL

In Subsection 9.1 we have classified the formulas that generate trivial self provers. Likewise we could study the question which formulas generate modalized self provers. In other words, for which φ do we have that $\varphi \wedge \Box\varphi$ is a modalized formula. In this subsection we take up this question for **GL**, make some remarks on it and pronounce a conjecture.

Definition 9.10. We say that *p occurs modalized in φ* if it occurs only under the scope of some modal operator. We say that *φ is modalized in p* if φ is equivalent to some φ' in which p occurs modalized. We say that φ is *strictly modalized* if every propositional variables occurs modalized in it. We call a formula φ *modalized* if it is equivalent to some strictly modalized formula.

We shall now give a model theoretic characterization of modalized formulae similar to our model theoretic characterization of Σ_1-formulae, Lemma 9.6.

Definition 9.11. Let T, r be a **GL**-model based on a rooted tree with root r. An *up-copy* of T, r is a rooted tree T', r which only (possibly) differs from T, r on the valuation of the propositional variables at r. We say that a formula φ is *preserved under up-copies* whenever $T, r \Vdash \varphi \Rightarrow T', r \Vdash \varphi$ holds for any rooted model T, r and any up-copy T', r of it.

Lemma 9.12. *(In* **GL***) φ is modalized iff. φ is preserved under up-copies.*

Proof. "$\Rightarrow$" This direction is easy. For example, one could use induction on the length of strictly modalized formulas employing Lemma 2.12.

"$\Leftarrow$" Suppose φ is not modalized. We can now write φ in a sort of disjunctive normal form, that is

$$\varphi \leftrightarrow \bigvee_i (l_i \wedge \varphi_i)$$

where l_i is a conjunction of literals and φ_i a strictly modalized formula. We claim that for some i, we have that $\varphi_i, \neg\varphi \nvdash \bot$. For, suppose $\forall i\ \varphi_i, \neg\varphi \vdash \bot$, then $\vdash \bigvee_i \varphi_i \to \varphi$, whence $\vdash \varphi \leftrightarrow \bigvee_i \varphi_i$. This is contrary to the assumption that φ is not modalized.

As **GL** is complete w.r.t. finite trees, we can find for the i for which $\nvdash \varphi_i \to \varphi$ a $T', x \Vdash \varphi_i \wedge \neg\varphi$. Consequently $T', x \Vdash \neg l_i$ and l_i is not empty. Thus, by just changing the valuation at x, we can get to an up-copy $T, x \Vdash l_i$. By the "$\Rightarrow$" part of this proof, we now know $T, x \Vdash \varphi_i$, whence $T, x \Vdash \varphi$. Clearly, T', x is an up-copy of T, x and φ is not preserved under up-copies. ⊣

Corollary 9.13. *"φ is modalized" is a decidable matter.*

Proof. As the proof of Corollary 9.7. ⊣

Again, the complexity is probably much lower than indicated by the above algorithm. Probably the following algorithm does also work if φ is given in "conjunctive normal form". Walk through φ and delete every modalized part. The remaining part should be provable or disprovable. By this algorithm we see that checking for modalizedness for φ in "CNF" is just $\mathsf{Co\text{-}NP}$-complete.

Conjecture 9.14. *$\varphi \wedge \Box\varphi$ is modalized iff.*
φ is modalized, or $\vdash \varphi \to \Diamond\top$

If φ is modalized, then clearly $\varphi \wedge \Box\varphi$ is modalized. On the other hand, if $\vdash \varphi \to \Diamond\top$, then $\vdash (\varphi \wedge \Box\varphi) \leftrightarrow \bot$ and is clearly modalized. The remaining implication is the harder implication. For example, it is not so hard to see that if $\nvdash \varphi \to \Diamond\top$ and $\varphi = p \wedge \psi$ or $\varphi = p \vee \psi$ with ψ a strictly modalized formula, that then $\varphi \wedge \Box\varphi$ is not modalized. Working with some sort of normal forms, the general case can probably be settled.

9.4 Essential Σ_1-ness and t.s.g.'s.

In Remark 7.7 we considered three different approaches for an alternative proof of the classification of the Σ_1-sentences of **IL**M. In this subsection we shall see that the second proposed approach may easily lead to the topic of t.s.g.'s. In that approach, we took up 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.

Lemma 9.15.

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 6.11.

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 $\vdash (\varphi \wedge \Box\neg\varphi) \leftrightarrow \Box\bot$). ⊣

Lemma 9.16.

$\Sigma(\varphi \wedge \Box\varphi) \wedge \Sigma(\varphi \wedge \Box\neg\varphi) \Rightarrow \Sigma(\varphi)$
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 9.15 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 $\vdash \varphi \wedge \Box\varphi \leftrightarrow \bigvee\!\!\!\bigvee_i \Box C_i$. In this case, clearly $\vdash \Box\bot \rightarrow \bigvee\!\!\!\bigvee_i \Box C_i \rightarrow \varphi$, whence, by Lemma 9.15 we conclude $\Sigma(\varphi \wedge \Box\neg\varphi)$. ⊣

By this lemma, we see that our original question (†) has been essentially reduced to a question on t.s.g.'s, a question that has proved to be at least as tough as the problem we originally started out with: a classification of the Σ_1-sentences of **IL**M.

Indeed, an answer to the question for which φ we have that $\Sigma(\varphi \wedge \Box\varphi) \Rightarrow \Sigma(\varphi)$ seems to require a thorough analysis of t.s.g.'s. We have a characterization of t.s.g.'s in **GL**. From this characterization it follows that (†) does certainly not hold for any formula φ, as we have seen that there are plenty of t.s.g.'s that are itself not Σ_1.

Thus, in general we may not assume that if $\neg\Sigma(\varphi)$, that then either $\neg\Sigma(\varphi\wedge\Box\varphi)$ or $\neg\Sigma(\varphi\wedge\Box\neg\varphi)$. Consequently, the $\Box\chi$ mentioned in the second approach of Remark 7.7 should in certain cases be sought elsewhere.

10 A variation: The logic Σ**IL**M

In this section we shall comment on some results obtained in a paper by Goris [Gor03]. In that paper, also a construction or step-by-step method was employed to obtain some modal completeness results.

This section contains no new results. However, we have included it to provide a richer context for this paper. We do formulate a new conjecture at the end of this section.

10.1 The logic Σ**IL**M

The language of Σ**IL**M is the modal language of interpretability with an additional unary modality Σ. The syntactical conventions concerning Σ are precisely the same as for the $\Box$-modality.

The additional Σ-modality is added to the language to capture the notion of "being equivalent to some Σ_1!-formula". Having this meaning in mind, the axioms of the logic Σ**IL**M come quite natural.

Definition 10.1. The logic Σ**IL**M is the smallest set of formulas being closed under Necessitation and Modus Ponens, containing **IL** (in the language with Σ) and all instantiations of the following axiom schemata.

$\Sigma\mathsf{1}$ $\Sigma\bot$

$\Sigma\mathsf{2}$ $\Sigma\Box A$

$\Sigma\mathsf{3}$ $\Sigma\Sigma A$

$\Sigma\mathsf{4}$ $\Sigma A\wedge\Sigma B\to\Sigma(A\wedge B)$

$\Sigma\mathsf{5}$ $\Sigma A\wedge\Sigma B\to\Sigma(A\vee B)$

$\Sigma\mathsf{6}$ $\Sigma A\wedge\Box(A\leftrightarrow B)\to\Sigma B$

$\Sigma\mathsf{7}$ $\Sigma A\to\Box\Sigma A$

$\Sigma\mathsf{8}$ $\Sigma A\to\Box(A\to\Box A)$

$\Sigma\mathsf{M}$ $\Sigma C\wedge(A\rhd B)\to A\wedge C\rhd B\wedge C$

A nice feature of Σ**IL**M is that it needs no new modal semantics. It can just be embedded in the semantics of **IL**M.

Definition 10.2. A Σ**IL**M-frame is an **IL**M-frame. The $\Vdash$-relation is extended to the new language by demanding that

$$w\Vdash\Sigma A\quad\text{iff.}\quad\forall u,v,w'[(w(R\cup S)^*w')\wedge(uS_{w'}v)\wedge(u\Vdash A)\Rightarrow v\Vdash A].$$

Lemma 10.3. Σ**IL**$\mathsf{M}\vdash A\Rightarrow F\models A$, *whenever* F *is a* Σ**IL**M*-frame.*

Proof. Straightforward. The **IL**M frame condition plays an essential role in $\Sigma\mathsf{2}$, $\Sigma\mathsf{3}$, and $\Sigma\mathsf{6}$. $\dashv$

An adequate labeled $\Sigma\mathbf{IL}\mathsf{M}$-frame is just an adequate $\Sigma\mathbf{IL}\mathsf{M}$-labeled $\mathbf{IL}\mathsf{M}$-frame for which moreover

$$uS_w v \wedge \Sigma A \in \nu(w) \text{ and } A \in \nu(u) \Rightarrow A \in \nu(v).$$

This condition should also be added to obtain the notion of a quasi-$\Sigma\mathbf{IL}\mathsf{M}$-frame. Lemma 6.5 immediately implies a closure lemma for quasi-$\Sigma\mathbf{IL}\mathsf{M}$-frames.

10.2 Modal and Arithmetical completeness

In Lemma 10.3 we have seen that $\Sigma\mathbf{IL}\mathsf{M}$ is sound w.r.t. $\Sigma\mathbf{IL}\mathsf{M}$-frames. In [Gor03] the logic is proved to be also complete.

Theorem 10.4. $\Sigma\mathbf{IL}\mathsf{M} \vdash \varphi$ *iff. for all* $\Sigma\mathbf{IL}\mathsf{M}$*-frames* F*, we have* $F \models \varphi$

The proof makes essentially use of the construction method, which is called the step-by-step method in [Gor03] and in [Joo98]. We find it instructive to mention this result in this paper, as it indicates a standard way to adopt the Main Lemma to modal languages with a richer or different signature.

In the logic $\Sigma\mathbf{IL}\mathsf{M}$ we have a new modality Σ. If we want a truth lemma to hold, also sentences containing Σ should be taken into account. As the quantifier complexity or the truth definition of the Σ-modality is $\forall$, no new deficiencies will arise. However, we should consider a new sort of problem.

Thus, a new sort of problem in a labeled frame F is a pair $\langle x, \neg\Sigma\varphi\rangle$ with $\neg\Sigma\varphi \in \nu(x)$, and such that for no y with $x(R\cup S)^* y$ and for no u, v with $uS_y v$ we have that $\varphi \in \nu(u)$ and $\neg\varphi \in \nu(v)$.

Such a problem can always be eliminated by taking $x = y$ and adding two labeled worlds with "$\varphi \in \nu(u) S_x \nu(v) \ni \neg\varphi$" in an adequate way to the labeled frame. That these labels $\varphi \in \nu(u) \subseteq_\Box \nu(v) \ni \neg\varphi$ can be found, can be proved by a lemma similar to Lemma 7.9.

The method of eliminating problems and deficiencies concerning the $\rhd$-modality can just be copied from the proof of Theorem 6.10. It is not hard to see that all the requirements on the Σ-modality just lift along with these elimination methods.

It is good to see that the Main Lemma can easily be adopted to modal logics with a different signature. Different modalities will yield different sort of problems and deficiencies. The distinction is always that a problem has an existential nature and a deficiency a universal one.

Arithmetical realizations are in a canonical way extended to formulas of $\Sigma\mathbf{IL}\mathsf{M}$. In [Gor03] an arithmetical completeness is proved for $\Sigma\mathbf{IL}\mathsf{M}$.

Theorem 10.5. $\Sigma\mathbf{IL}\mathsf{M} \vdash \varphi \quad$ *iff.* $\quad \forall * \ T \vdash \varphi^*$

On purpose we did not specify the T in the theorem, as there is some freedom. Either T is essentially reflexive, or T is an extension of $\mathsf{I}\Sigma_1$. In the latter case $\rhd$ should be translated to a formalization of Π_1-conservativity rather than interpretability.

A nice observation that we can draw from Theorem 10.5 is that Lemma 7.3 can actually be reversed. Thus,

$$\begin{array}{c} \forall * \ T \vdash \Sigma(\varphi^*) \\ \text{iff.} \\ \forall \alpha, \beta \forall * \ T \vdash \alpha \rhd \beta \to \alpha \wedge \varphi^* \rhd \beta \wedge \varphi^*. \end{array}$$

Here T is of course an essentially reflexive theory. We conjecture that we have a far more general arithmetical fact. Namely, that for any arithmetical formula γ (thus not necessarily the translation of a modal formula) we have that

$$\begin{array}{c} T \vdash \Sigma(\gamma) \\ \text{iff.} \\ \forall \alpha, \beta \ T \vdash \alpha \rhd \beta \to \alpha \wedge \gamma \rhd \beta \wedge \gamma. \end{array}$$

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

To start, let us recall the schema M_0.

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

Let us also recall that $\mathbf{IL}\mathsf{M}_0$ is the logic we obtain when adding this schema to **IL**. This section is devoted to showing the following theorem.[23]

Theorem 11.1. $\mathbf{IL}\mathsf{M}_0$ *is a complete logic.*

In the light of Remark 4.22 a proof of Theorem 11.1 boils down to giving the four ingredients mentioned there. Section 11.3, 11.4, 11.5 and 11.6 below contain those ingredients. Before these main sections we have in Section 11.2 some preliminaries but we start in Section 11.1 with an overview of the difficulties we encounter during the application of the construction method to $\mathbf{IL}\mathsf{M}_0$. We round things up in Section 11.7 and finish in Section 11.8 with some reflections on what we have done.

11.1 Overview of difficulties

We shortly review the main construction method and see what difficulties we might encounter when we apply it to $\mathbf{IL}\mathsf{M}_0$.

Roughly the construction proceeds as follows. We identify a problem or a deficiency in a labeled frame F. We solve this problem or deficiency by extending F to a labeled frame F'. Then we extend F' to a frame F'' in which all the frame conditions for the logic under consideration hold. We repeat this procedure until no problems nor deficiencies are present any more.

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 $\nu(w) \prec_B \nu(y)$.

[23] A proof of this theorem was first given in [Joo98]. In this section we fill in some missing details.

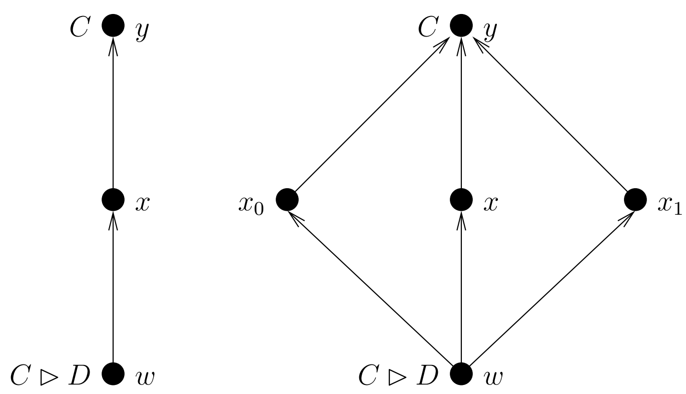

Figure 3: A deficiency in w w.r.t. y

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 requires that we also have xRz. So, from 2. and the M_0–frame condition we obtain

$$wRxRyS_wy'Rz \to \nu(x) \prec \nu(z).$$

If we put $\Delta \subseteq_\Box \Delta' :\Leftrightarrow \{\Box A \mid \Box A \in \Delta\} \subseteq \Delta'$ then a sufficient (and in certain sense necessary) condition is,

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

Let us illustrate some difficulties concerning this condition by some examples. Consider the left 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 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 the lemma below. We omit a proof, but one 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. *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 $\Delta \subseteq_\Box \Delta'$.*

But now look at the right model in Figure 3. We have at least for two different worlds x, say x_0 and x_1, that $wRxRy$. The above lemma 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 handled this difficulty by ensuring that it is enough to consider only one of the worlds in between w and y. To be precise, we ensured $\nu(x') \subseteq_\Box \nu(x)$ or $\nu(x) \subseteq_\Box \nu(x')$.

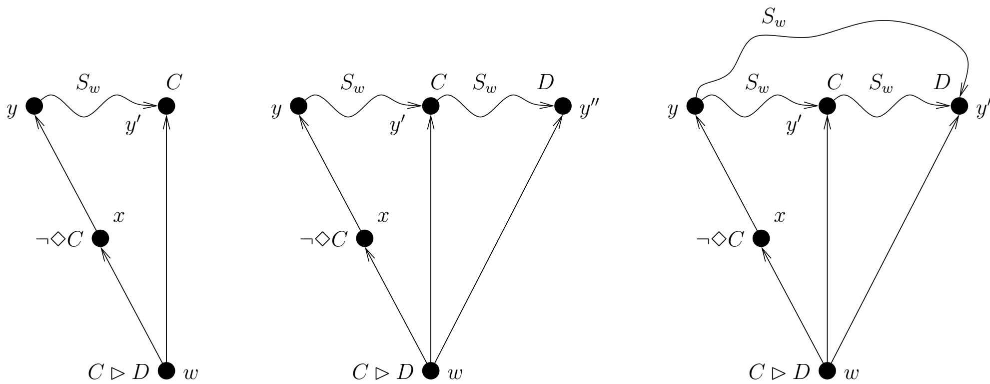

Figure 4: A deficiency in w w.r.t. y'

But now some difficulties concerning Item 1. occur. In the situations in Figure 3 we were asked to solve a deficiency in w w.r.t. y. But we actually solved one in w w.r.t. some x in between w and y. We can let y' solve both deficiencies if we have that $\nu(w) \prec_A \nu(y')$ whenever $y \in \mathcal{C}_w^A$. And this is assured by ensuring that $wRxRy \in \mathcal{C}_w^A$ implies $\nu(w) \prec_A \nu(x)$.

We are not there yet. Consider the leftmost model in Figure 4. 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 above lemma is not applicable. Below we formulate and proof a more complicated version of the above lemma 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 that lemma ensures us that we can, indeed, choose $\nu(y')$ appropriate.

11.2 Preliminaries

Definition 11.2 (T^{tr}, T^*, $T;T'$, T^1, $T^{\geq 2}$, $T \cup T'$). Let T and T' be binary relations on a set W. We introduce the following notations.

1. T^{tr} is the transitive closure of T.
2. T^* is the transitive reflexive closure of T.
3. $xT;T'y \Leftrightarrow \exists t\ xTtT'y$
4. $xT^1y \Leftrightarrow xTy \wedge \neg\exists t\ xTtTy$
5. $xT^{\geq 2}y \Leftrightarrow xTy \wedge \neg(xT^1y)$
6. $xT \cup T'y \Leftrightarrow xTy \vee xT'y$

Definition 11.3 ($\mathcal{S}_w$). 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.

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

Definition 11.4 (adequate $\mathbf{IL}\mathsf{M}_0$–frame). Let $F = \langle W, R, S, \nu\rangle$ be an adequate frame. We say that F is an adequate $\mathbf{IL}\mathsf{M}_0$–frame iff. the following additional properties hold.[24]

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

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 11.9 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 11.5 ($K(F)$, 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 11.6. *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 $K(F_0) = K(F_1)$.*

Proof. Put

$$K_0 = K(F_0)$$
$$K_1 = K(F_1)$$

We show $K_0 \subseteq K_1$. By symmetry the lemma then follows. We show that K_1 satisfies properties 1., 2. and 3. of the definition of K_0 (Definition 11.5).

1. By assumption, $R_0 \subseteq K_1$.
2. By definition of K_1, $K_1 = (K_1)^{\mathrm{tr}}$.
3. Suppose $wK_1xK_1^1y(\mathcal{S}_w)^{\mathrm{tr}}y'K_1^1z$. Then, by definition of K_1, xK_1z.

So, since K_0 is the smallest binary relation that satisfies all of these properties we conclude $K_0 \subseteq K_1$. ⊣

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.

[24]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.

Lemma 11.7 (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), we will show that T^{tr} satisfies the three properties of the definition of K (Definition 11.5).

$R \subseteq T^{\text{tr}}$ and the transitivity of T^{tr} are clear. So suppose $wT^{\text{tr}}x(T^{\text{tr}})^1y(\mathcal{S}_w)^{\text{tr}}y'(T^{\text{tr}})^1z$. Since $(T^{\text{tr}})^1 = T^1$ we have $wT^{\text{tr}}xT^1y(\mathcal{S}_w)^{\text{tr}}y'T^1z$. By assumption 2. on T we obtain $xT^{\text{tr}}z$. We have shown (a).

To show (b) assume xK^1y. Then by (a). $xT^{\text{tr}}y$. But if $xT^{\geq 2}y$ then, since $T \subseteq K$, also $xK^{\geq 2}y$. ⊣

Another entity that changes during the construction of an $\mathbf{IL}\mathsf{M}_0$–frame out of a quasi–frame is the critical cone (which is defined for all logics $\mathbf{IL}\mathsf{X}$, see Definition 4.6). In accordance with the above definition of $K(F)$, we also like to predict what eventually becomes the critical cone.

Definition 11.8 ($\mathcal{N}^C_w$). For any quasi–frame F we define $\mathcal{N}^C_w$ to be the smallest set such that

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

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 $\mathbf{IL}$–frameconditions, but under all the $\mathbf{IL}\mathsf{M}_0$–frameconditions.

Definition 11.9 (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. $x \in \mathcal{N}^A_w \to \nu(w) \prec_A \nu(x)$
9. $wKxKy(S_w \cup K)^*y' \to \nu(x) \subseteq_\Box \nu(y')$
10. $xS_wy \to x(\mathcal{S}_w \cup R)^*y$
11. $wKxK^1y(\mathcal{S}_w)^{\text{tr}}y'K^1z \to x(K^1)^{\text{tr}}z$
12. $xRy \to x(R^1)^{\text{tr}}y$

Lemma 11.10. *If F is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame, then $K = (K^1)^{tr}$.*

Proof. It is enough to show that K^1 satisfies the three conditions on T in Lemma 11.7. Item 2. is clear by Property 11. of quasi–$\mathbf{IL}\mathsf{M}_0$–frames. To see Item 1. we note that $R \subseteq K$ implies $(R^1)^{\text{tr}} \subseteq (K^1)^{\text{tr}}$. So, $R \subseteq (R^1)^{\text{tr}}$ implies $R \subseteq (K^1)^{\text{tr}}$. ⊣

Lemma 11.11. *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 11.5 and additionaly we have the following.*

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

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. We clearly have that K', K'' and K''' satisfies the three defining properties of K. So, we have $K \subseteq K'$, $K \subseteq K''$, $K \subseteq K'''$. So we are done when we have seen that K satisfies $3'$., $3''$., $3'''$. We can see $3'$. with induction on the minimal number of K steps in the $(\mathcal{S}_w \cup K)$–path from y to y'. Item $3''$. follows by Lemma 11.10. Item $3'''$. follows by combining $3'$. and $3''$. ⊣

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 11.12 ($\mathbf{IL}\mathsf{M}_0$–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 two differences are that we add a fifth entry to the list of imperfections.

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

And we work with an enumeration on

$$C' := C \cup (\{4\} \times W^5).$$

Where C is defined as in Lemma 5.2, nameley

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

First we will show 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. This implies that $\hat{F}$ is an adequate $\mathbf{IL}\mathsf{M}_0$–frame.

To see that each F_n is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame, we proceed with induction on n. The case $n = 0$ holds by assumption on $F = F_0$. So, assume that F_n is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. Let for each $n \geq 0$

$$
\begin{aligned}
K^n &= K(F_n)\\
R^n &= R^{F_n}\\
S^n &= S^{F_n}
\end{aligned}
$$

Claim 11.12a. 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 show Claim (a) and (b) simultainously. Item (c) is an immediate corollary to (b). We distinguish cases according to which imperfection is dealt with in the step from F_n to F_{n+1}.

(i) Suppose the imperfection under consideration concerns the transitivity of R^n. First we show (a).

Suppose $xR^{n+1}y$. If xR^ny then we are done. So suppose this is not so. Then for some t, xR^ntR^ny. So, xK^ntK^ny. By transitivity of K^n we conclude xK^ny.

(b) follows from (a) and the fact that $S^n = S^{n+1}$.

(ii)-(iv) Suppose the imperfection under consideration concerns the reflexivity of S^n, the transitivity of S^n or the inclusion of S^n in R^n. Both (a) and (b) are clear in these cases.

(v) Suppose the imperfection under consideration concerns the M_0 frame condition.

We first show (a).

Suppose $xR^{n+1}y$. If xR^ny then we are done. So, suppose this is not so. Then for some u, v, v' we have $uR^nxR^nvS_u^nv'R^ny$. So, $uK^nxK^nvS_u^nv'K^ny$ and thus by Lemma 11.11 we conclude xK^ny.

(b) follows from (a) and the fact that $S^{n+1} = S^n$.

$\dashv$

Now let us show that F_{n+1} is a quasi–frame. We only have to consider the case in which F_{n+1} is constructed out of F_n by solving an imperfection of the form (v). The other cases can be copied from the proof of Lemma 5.2.

1. R^{n+1} is conversly well–founded.

   By Claim 11.12a-(a), any R^{n+1} chain is a K^n. And K^n is conversely well–founded.

2. $xS_w^{n+1}y \to wR^{n+1}x, y$.

   This is clear, since $S^{n+1} = S^n$ and $R^n \subseteq R^{n+1}$.

3. $xR^{n+1}y \to \nu(x) \prec \nu(y)$.

   Suppose $xR^{n+1}y$. By Claim 11.12a-(b), we have xK^ny and thus, since F_n is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame, $\nu(x) \prec \nu(y)$.

4. $A \neq B \Rightarrow F_{n+1} \models \mathcal{G}^A_w \cap \mathcal{G}^B_w = \emptyset$.
   It is enough to show, that for all formulas C,
   $$\forall wy\ F_{n+1} \models y \in \mathcal{G}^C_w \Rightarrow F_n \models y \in \mathcal{G}^C_w.$$
   Suppose $F_{n+1} \models y \in \mathcal{G}^A_w$. Then we have for some $k \geq 1$ and x_1 such that $\nu^{n+1}(w, x_1) = A$,
   $$wR^{n+1}x_1(R^{n+1} \cup S^{n+1}) \cdots (R^{n+1} \cup S^{n+1})x_k = y.$$
   If, for some $i < k$, not $x_i(R^n \cup S^n)x_{i+1}$ then, for some u, v and v', $uR^n x_i R^n v S^n_u v' R^n x_{i+1}$. So, in any case, $F_n \models y \in \mathcal{G}^A_w$.
5. $F_{n+1} \models x \in \mathcal{C}^A_w \Rightarrow \nu(w) \prec_A \nu(x)$.
   Suppose $F_{n+1} \models x \in \mathcal{C}^A_w$. Then $F_n \models x \in \mathcal{N}^A_w$, by Claim 11.12a-(c). And thus $\nu(w) \prec_A \nu(x)$.

Now we know that F_{n+1} is a quasi–frame and thus $K(F_{n+1})$ is defined. Before we show that F_{n+1} is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame we strengthen Claim 11.12a.

Claim 11.12b. For all w, x and A we have the following.

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

Proof. Item 1. follows by Claim 11.12a and Lemma 11.6. Item 2. follows from Item 1. and Claim 11.12a-(b). Item 3. is an immediate corollary of item 2.. $\dashv$

Now we show that F_{n+1} is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. So we run through the properties of Definition 11.9.

6. K^{n+1} is conversly well–founded.
   By Claim 11.12b-1.
7. $xK^{n+1}y \to \nu(x) \prec \nu(y)$.
   By Claim 11.12b-1.
8. $F_{n+1} \models x \in \mathcal{N}^A_w \Rightarrow \nu(w) \prec_A \nu(x)$.
   By Claim 11.12b-3.
9. $wK^{n+1}xK^{n+1}y(S^{n+1}_w \cup K^{n+1})^* y' \to \nu(x) \subseteq_\Box \nu(y')$.
   By Claim 11.12b-1.-2.
10. $xS^{n+1}_w y \to x(\mathcal{S}_w \cup R^{n+1})^* y$.
    Immediate.
11. $wK^{n+1}xK^{n+1^1}y(\mathcal{S}^{n+1}_w)^{\mathrm{tr}}y'K^{n+1^1}z \to x(K^{n+1^1})^{\mathrm{tr}}z$.
    By Claim 11.12b-1.
12. $xR^{n+1}y \to x(R^{n+1^1})^{\mathrm{tr}}y$.
    Since F is a qausi we have $K = (K^1)^{\mathrm{tr}}$. So, since $xR^{n+1}y$ implies xKy we have $xR^{n+1}y$ implies $x(K^1)^{\mathrm{tr}}y$. But $R^{n+1} \subseteq K$. Thus we conclude $x(R^{n+1^1})^{\mathrm{tr}}y$.

So we have shown that each F_n is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. We put

$$\hat{F} = \bigcup_{n \geq 0} F_n.$$

We have to show that $\hat{F}$ is an adequate $\mathbf{IL}\mathsf{M}_0$–frame. To show that $\hat{F}$ is an adequate frame only the conversely well–foundedness of R needs some attention. The other properties lift along with the proof given for Lemma 5.2. By Claim 11.12b-1. any R-chain in $\hat{F}$ is a K-chain in F. And since F is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame, such a chain is finite.

To show that $\hat{F}$ is an adequate $\mathbf{IL}\mathsf{M}_0$–frame we have to check the four additional properties of Definition 11.4. Let $\langle \hat{W}, \hat{R}, \hat{S}, \hat{\nu} \rangle = \hat{F}$.

4. $w\hat{R}x\hat{R}y\hat{S}_w y'\hat{R}z \rightarrow x\hat{R}z$
   In the terminology of Lemma 5.2, an example for failure of this property is an imperfection of type (v). A proof that $\hat{F}$ does not have imperfections of this type is completely similair to the proof that $\hat{F}$ does not have imperfections of types (i)–(iv), as given for Lemma 5.2,
5. $w\hat{R}x\hat{R}y\hat{S}_w y' \rightarrow \hat{\nu}(x) \subseteq_\Box \hat{\nu}(y')$
   Suppose $w\hat{R}x\hat{R}y\hat{S}_w y'$. Then, for some $n \geq 0$, $wR^n xR^n yS^n_w y'$. Thus, since F_n is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame, $\hat{\nu}(x) \subseteq_\Box \hat{\nu}(y')$.
6. $x\hat{S}_w y \rightarrow x(\hat{\mathcal{S}}_w \cup \hat{R})^* y$
   If $x\hat{S}_w y$ then, for some $n \geq 0$, $xS^n_w y$. So, since F_n is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame, $x(\mathcal{S}_w \cup R^n)^* y$. Since $\hat{\mathcal{S}}_w = \mathcal{S}_w$ we obtain $x(\hat{\mathcal{S}}_w \cup \hat{R})^* y$.
7. $x\hat{R}y \rightarrow x(\hat{R}^1)^{\mathrm{tr}} y$
   $\hat{R} = K$, so this follows immediate from the fact that $K = (K^1)^{\mathrm{tr}}$.

⊣

Lemma 11.13. *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. Let

$$T = \{(x, y) \mid \exists z\,(\nu(x) \subseteq_\Box \nu(z) \wedge x(R \cup S)^* zRy)\}.$$

Claim 11.13a. T is transitive.

Proof. Suppose $xTyTz$. Then for some t_x with $\nu(x) \subseteq_\Box \nu(t_x)$ and some t_y with $\nu(y) \subseteq_\Box \nu(t_y)$

$$x(R \cup S)^* t_x Ry(R \cup S)^* t_y Rz.$$

Also $\nu(t_x) \prec \nu(y)$. So

$$\nu(x) \subseteq_\Box \nu(t_y).$$

Also clearly

$$x(R \cup S)^* t_y Rz.$$

Thus we conclude xTz. ⊣

Claim 11.13b. $\{(x,y) \mid \exists t\,(\nu(x) \subseteq_{\Box} \nu(t) \wedge xT;(S\cup K)^{*}tTy)\} \subseteq T$.

Proof. Suppose that for some t with $\nu(x) \subseteq_{\Box} \nu(t)$,

$$xT;(S\cup K)^{*}tTy.$$

Then for some t' with $\nu(t) \subseteq_{\Box} \nu(t')$,

$$xT;(S\cup K)^{*}t(R\cup S)^{*}t'Ry.$$

Now $\nu(x) \subseteq_{\Box} \nu(t')$ and $x(R\cup S)^{*}t'Ry$. So we conclude xTy. ⊣

We define

$$K' = K \cap T.$$

We have to show that $K' = K$. As $K' \subseteq K$ is trivial, we will show $K \subseteq K'$. We show that K' satisfies properties 1., 2. and 3. of Definition 11.5.

1. $R \subseteq K'$, since $R \subseteq T$ and $R \subseteq K$x.
2. If $xK'yK'z$ then $xKyKz$ and thus xKz. Also, $xTyTz$. So, by Claim 11.13a, xTz.
3. If $wK'xK'y(S_w \cup K)^{*}y'K'z$ then $wKxKy(S_w\cup K)^{*}y'Kz$. So, xKz and, since F is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame, $\nu(x) \subseteq_{\Box} \nu(y')$. Also, $xTy(S_w \cup K)^{*}y'Tz$. Thus, by Claim 11.13b, xTz.

Since K is the smallest binary relation that satisfies these properties we conclude $K \subseteq K'$. ⊣

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 11.14. *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. We define

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

Clearly K^- satisfies the properties 1., 2. and 3. of the definition of K_G (Definition 11.5). Thus, since K_G is the smallest such relation,

$$K_G \subseteq K^-.$$

In what follows we will show

$$K^- \subseteq K_G. \tag{9}$$

Or, to put it differently, $K^- - K_G = \emptyset$. For this, consider

$$K' = K - (K^- - K_G).$$

We will show that

$$K \subseteq K'. \tag{10}$$

Since $K^- - K_G \subseteq K$, this implies $K^- - K_G = \emptyset$. (And thus we can conclude (9).) We show that K' satisfies properties 1., 2. and 3. of the definition of K (Definition 11.5).

1. By definition of K we have

$$R \subseteq K. \tag{11}$$

Next we will show

$$R \cap (K^- - K_G) = \emptyset. \tag{12}$$

It is enough to show $R \cap K^- \subseteq K_G$. Suppose xRy and xK^-y. Then $xR\restriction_{W^-}y$ and thus by assumption (b) of this lemma, xK_Gy. Thus we have shown (12). It is easy to see that (11) and (12) together imply

$$R \subseteq K'.$$

2. Suppose $xK'yK'z$. Then in particular $xKyKz$. And thus

$$xKz. \tag{13}$$

First, suppose $x, y, z \in W^-$. Since K^- is simply the restriction of K to W^- we have xK^-yK^-z. So, by definition of K', xK_GyK_Gz. By transitivity of K_G we obtain

$$xK_Gz. \tag{14}$$

Combining (13) and (14) we get $xK'z$.

Next suppose some of x, y, z are members of $W - W^-$. By assumption (a) of this lemma and by Lemma 11.13 we then have that certanly $z \in W - W^-$. Since K^- is the restriction of K to W^-, this implies

$$\text{not } xK^-z. \tag{15}$$

Combining (13) and (15) we get $xK'z$.

3. Suppose $wK'xK'y(\mathcal{S}_w)^{\mathrm{tr}}y'K'z$. Then $wKxKy(\mathcal{S}_w)^{\mathrm{tr}}y'Kz$. And thus

$$xKz. \tag{16}$$

First suppose $w, x, y, y', z \in W^-$. Then $wK^-xK^-y(\mathcal{S}_w)^{\text{tr}}y'K^-z$. And thus by defintin of K', $wK_GxK_Gy(\mathcal{S}_w)^{\text{tr}}y'K_Gz$. This gives

$$xK_Gz. \tag{17}$$

Combining (16) and (17) we get $xK'z$.

Next suppose some of w, x, y, y', z are members of $W - W^-$. Then by assumption (a) of this lemma and by Lemma 11.13, $z \in W - W^-$. Since K^- is the restriction of K to W^-, this implies

$$\text{not } xK^-z. \tag{18}$$

Combining (16) and (18) we get $xK'z$.

Since K is the smallest binary relation that satisfies these properties, we conclude (10).

$\dashv$

We finish the basic preliminaries with a somewhat complicated variation of Lemma 4.21.

Lemma 11.15. *Let* Γ *and* Δ *be* MCS*'s.* $\Gamma \prec_C \Delta$.

$$P \rhd Q, S_1 \rhd T_1, \ldots, S_n \rhd T_n \in \Gamma$$

and

$$\Diamond P \in \Delta.$$

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 *(e.g.* $\Delta \subseteq_\Box \Delta_i$*),*
- $Q \in \Delta_0$*,*
- *For all* $1 \leq j \leq n$*,* $S_j \in \Delta_h \Rightarrow$ *for some* $i \leq k$*,* $T_j \in \Delta_i$*.*

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

$$A :\Leftrightarrow \bigwedge \mathcal{A},$$

$$B :\Leftrightarrow \bigvee \mathcal{B}.$$

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

$$Q, \overline{S}_I, \Box A, \neg B \wedge \Box\neg B \vdash \bot \tag{19}$$

or,

$$\text{for some } i \notin I, \quad T_i, \overline{S}_I, \Box A, \neg B \wedge \Box\neg B \vdash \bot. \tag{20}$$

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

$$T_{i_k}, \overline{S}_{I_k}, \Box A, \neg B \wedge \Box\neg B \vdash \bot. \tag{21}$$

Additionally, we will verify that for each k

$$\text{(19) does not hold with } I_k \text{ for } I.$$

We will define i_k with induction on k. We define $I_1 = \emptyset$. And by Lemma 4.21, (19) does not hold with $I = \emptyset$. Moreover, because of this, (20) must be true with $I = \emptyset$. So, there exists some $i \in \{1, \dots, n\}$ 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

$$Q, \overline{S}_{I_{k+1}}, \Box A, \neg B \wedge \Box\neg B \nvdash \bot. \tag{22}$$

Let us suppose that this is not so. Then

$$\vdash \Box(Q \to \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \dots \vee S_{i_k}). \tag{23}$$

So,

$$
\begin{array}{lll}
\Gamma \vdash P \rhd Q & & \\
& \rhd\, \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \dots \vee S_{i_{k-1}} \vee S_{i_k} & \text{by (23)} \\
& \rhd\, \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \dots \vee S_{i_{k-1}} \vee T_{i_k} & \\
& \rhd\, \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \dots \vee S_{i_{k-1}} \vee (T_{i_k} \wedge \Box A \wedge \neg B \wedge \Box\neg B \wedge \overline{S}_{I_k}) & \\
& \rhd\, \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \dots \vee S_{i_{k-1}} & \text{by (21)} \\
& \quad \vdots & \\
& \rhd\, \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} & \\
& \rhd\, \Diamond\neg A \vee B \vee \Diamond B \vee T_{i_1} & \\
& \rhd\, \Diamond\neg A \vee B \vee \Diamond B \vee (T_{i_1} \wedge \Box A \wedge \neg B \wedge \Box\neg B) & \\
& \rhd\, \Diamond\neg A \vee B \vee \Diamond B. & \text{by (21), with } k = 1.
\end{array}
$$

So by 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 4.21 there exists some MCS Δ with $\Gamma \prec_C \Delta$ that contains $B \vee \Diamond B$. This is a contradiction, so we have shown (22).

But now, since (22) is indeed true, and thus (19) with I_{k+1} for I is false, (20) 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 $I = \{1, 2, \ldots, n\}$, (20) cannot be true. Thus, for $I = \{1, 2, \ldots, n\}$, (19) 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 (Lemma 4.19) applied for $\mathbf{IL}\mathsf{M}_0$, we need one more definition and a lemma.

Definition 11.16 ($\subset_1$, $\subset$). 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 11.17. *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 $x \subset y \Rightarrow \nu(x) \subseteq_\Box \nu(y)$.*

11.3 Frame condition

Theorem 11.18. *For an $\mathbf{IL}$-frame $F = \langle W, R, S, \nu\rangle$ we have*

$$\forall wxyy'z\ (wRxRyS_wy'Rz \to xRz) \Leftrightarrow F \models \mathsf{M}_0.$$

Proof. Suppose F is an $\mathbf{IL}\mathsf{M}_0$–frame, let $\overline{F}$ be a model based on F and $w \in W$. Suppose $w \Vdash A \rhd B$. Pick any world $x \in W$ for which wRx and $x \Vdash \Diamond A \wedge \Box C$. Then there exists some world y, such that xRy and $y \Vdash A$. R is transitive, so there exists some y', with yS_wy' and $y' \Vdash B$. $xRyS_wy'$ implies xS_wy', so we are done once we have shown that $y' \Vdash \Box C$. Let us assume, for a contradiction, that this is not so. Then there exists some z with $y'Rz$ and $z \Vdash \neg C$. Since F is an $\mathbf{IL}\mathsf{M}_0$–frame we have xRz. So, in particular $x \Vdash \Diamond\neg C$. A contradiction.

Suppose $F \models \mathsf{M}_0$. Choose $w, x, y, y', z \in W$ such that $wRxRyS_wy'Rz$. We have to show that xRz. Let p, q and s be distinct proposition variables. Define a model $\overline{F}$, based on F as follows.

$$v \Vdash p \Leftrightarrow v = y$$
$$v \Vdash q \Leftrightarrow v = y'$$
$$v \Vdash s \Leftrightarrow xRv$$

Since $F \models \mathsf{M}_0$, in particular we have $w \Vdash p \rhd q \to \Diamond p \wedge \Box s \rhd q \wedge \Box s$. Also, by definition of $\Vdash$, $w \Vdash p \rhd q$ and $x \Vdash \Diamond p \wedge \Box s$. So there must exists some world t for which $t \Vdash q \wedge \Box s$. The only candidate for this is y'. And thus we must have $z \Vdash s$. By definition of $\Vdash$ this give xRz. $\dashv$

11.4 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 $\subseteq_\Box$

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

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

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

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

$\mathcal{I}_{\mathsf{M}_0}$ 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.[25]

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

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

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

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

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

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

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

$\mathcal{J}_{\nu_4}$ If $x(\mathcal{S}_w)^{\mathrm{tr}}y$, 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 the first subsection.

$\mathcal{I}_{\mathcal{D}}$ ensures the boundedness of all the frames. Which is needed to ensure the conversely well–foundedness of R in the end.

$\mathcal{I}_{\mathsf{M}_0}$ 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 11.19. *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.*

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

Proof. Only $\mathcal{I}_{\mathcal{D}}$ and $\mathcal{I}_{\mathrm{d}}$ needs 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 11.12).

We first tread $\mathcal{I}_{\mathcal{D}}$. We only have to consider the case in which F_{n+1} is constructed out of F_n by solving an imperfection of the form (v). So, we have $wR^{F_n}xR^{F_n}yS_w^{F_n}y'R^{F_n}z$. Since F_n is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame we have $\nu(x) \subseteq_{\Box} \nu(y')$. And $\exists\, A{\in}(\nu(z)\setminus\nu(y'))\cap\{\Box D \mid D \in \mathcal{D}\}$. So, this very same A cannot be an element of $\nu(x)$. Thus $\exists\, A{\in}(\nu(z)\setminus\nu(x))\cap\{\Box D \mid D \in \mathcal{D}\}$.

Now let us tread $\mathcal{I}_{\mathrm{d}}$. We might get in trouble if at some point in the construction we add wRx. But the premise of $\mathcal{I}_{\mathrm{d}}$ implies wRx and since the premise is stable we see that $\mathcal{I}_{\mathrm{d}}$ is preserved.

$\dashv$

Lemma 11.20. *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.*

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 11.21. *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 11.22. *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*

$$x\dagger_v,\ y \in \mathcal{N}_w^A \Rightarrow x \in \mathcal{N}_w^A.$$

Proof. Suppose $x\dagger_v$ and $y \in \mathcal{N}_w^A$. Then, by Corollary 11.21, $\nu(w,y) = A$ 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 11.17–1., tKx. $\dashv$

11.5 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. Fix some $\mathbf{b} \notin W$. Using Lemma 4.20 we find a MCS $\Delta_{\mathbf{b}}$, such that $\nu(\mathbf{a}) \prec_B \Delta_{\mathbf{b}}$ and $A, \Box\neg A \in \Delta_{\mathbf{b}}$. Put

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

The frames F and $\hat{F}$ satisfy the conditions of Lemma 11.14. Thus we have

$$\forall xy{\in}F\ xKy \Leftrightarrow x\hat{K}y. \tag{24}$$

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 $\hat{K}$ resplaced by K. In what follows we will use this extensively. With and without mention.

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

Proof. We run through the properties (1.–5.) of Definition 5.1 (quasi–frames) and the properties (6.–10.) of Definition 11.9 (quasi–$\mathbf{IL}\mathsf{M}_0$–frames) and the remaining ones in Definition 5.1 (quasi–frames).

1. $\hat{R}$ is conversely well-founded

   This is clear.
2. $x\hat{S}_w y \rightarrow w\hat{R}x, y$.

   Evident, since $\hat{S} = S$ and if $x\hat{S}_w y$ then $w, x, y \in F$.
3. $w\hat{R}x \rightarrow \hat{\nu}(w) \prec \hat{\nu}(x)$

   We only have to consider the case $w = \mathbf{a}$ and $x = \mathbf{b}$. This case is clear by choice of $\Delta_{\mathbf{b}}$.
4. $C \neq D \Rightarrow \hat{F} \models \mathcal{G}^C_w \cap \mathcal{G}^D_w = \emptyset$.

   For all $x \in F$, $\hat{F} \models x \in \mathcal{G}^C_w \Leftrightarrow F \models x \in \mathcal{G}^C_w$. So we only have to consider the case $\hat{F} \models \mathbf{b} \in \mathcal{G}^C_w$. If $w = \mathbf{a}$ then we are done. So suppose $w \neq \mathbf{a}$. Then there exists some $y \in F$ such that $\nu(w, y) = C$ and $y(\hat{R} \cup \hat{S})^* x$. Similarly, if $\hat{F} \models x \in \mathcal{G}^D_w$ then for some $y' \in F$ with $\nu(w, y') = D$ we have $y''(\hat{R} \cup \hat{S})^* x$. But then $y(\hat{R} \cup \hat{S})^* \mathbf{a}$ and $y'(\hat{R} \cup \hat{S})^* \mathbf{a}$. Which implies $F \models \mathbf{a} \in \mathcal{G}^D_w \cap \mathcal{G}^C_w$. Conclusion: $D = C$.
5. $\hat{F} \models x \in \mathcal{C}^A_w \Rightarrow \hat{\nu}(w) \prec_A \hat{\nu}(x)$

   If $w \neq \mathbf{a}$ then this follows since for all r, s, t, C, $\hat{\nu}(r) \prec_C \hat{\nu}(s) \prec \hat{\nu}(t)$ implies $\hat{\nu}(r) \prec_C \hat{\nu}(t)$. If $w = \mathbf{a}$ then this follows by the choice of $\Delta_{\mathbf{b}}$.
6. $\hat{K}$ is conversely well-founded.

   By (24) any $\hat{K}$-chain that is not a K-chain must include $\mathbf{b}$. But any such chain is obviously finite.
7. $x\hat{K}y \rightarrow \hat{\nu}(x) \prec \hat{\nu}(y)$.

   If $x, y \in F$, then this follows from (24). If one of x and y are in $\hat{F} - F$ then it is clear that this can only be y. So assume $y = \mathbf{b}$. By Lemma 11.13 we have for some z,

   $$\nu(x) \subseteq_{\Box} \nu(z)$$
   $$x(\hat{R} \cup \hat{S})^* z\hat{R}\mathbf{b}.$$

   The only candidate for such a z is $\mathbf{a}$. We have chosen $\hat{\nu}(\mathbf{b})(= \Delta_{\mathbf{b}})$ such that

   $$\hat{\nu}(\mathbf{a}) \prec \hat{\nu}(\mathbf{b}).$$

   So, we conclude $\hat{\nu}(x) \prec \hat{\nu}(\mathbf{b})$.
8. $\hat{F} \models x \in \mathcal{N}^C_w \Rightarrow \hat{\nu}(w) \prec_C \hat{\nu}(x)$.

   We have

   $$\forall x w {\in} F \; F \models x \in \mathcal{N}^C_w \Leftrightarrow \hat{F} \models x \in \mathcal{N}^C_w.$$

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 11.22, we have for some $x \in F$, $F \models x \in \mathcal{N}_w^C$ and $x\hat{K}\mathbf{b}$. By item 7. this implies $\hat{\nu}(x) \prec \hat{\nu}(\mathbf{b})$. So since $\hat{\nu}(w) \prec_C \hat{\nu}(x)$ we have $\hat{\nu}(w) \prec_C \hat{\nu}(\mathbf{b})$.

9. $\forall wxyy' \, (w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* y' \rightarrow \hat{\nu}(x) \subseteq_\Box \hat{\nu}(y'))$.
   Let $w, x, y, y' \in \hat{F}$ and suppose $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* y'$. We can assume that $y \neq y'$.
   Case 1: All of w, x, y, y' are in F. Then from (24) it follows that
   $$wKxKy(S_w \cup K)^* y'.$$
   So, $\nu(x) \subseteq_\Box \nu(y')$. And thus $\hat{\nu}(x) \subseteq_\Box \hat{\nu}(y')$.
   Case 2: Some of w, x, y, y' are in $\hat{F} - F$. Then clearly $y' = \mathbf{b}$ and, for some v, $wKxKy(S_w \cup K)^* vK\mathbf{b}$. Now $\nu(x) \subseteq_\Box \nu(v)$. And since $v\hat{K}\mathbf{b}$ implies $\hat{\nu}(v) \prec \hat{\nu}(\mathbf{b})$ we certainly have $\hat{\nu}(x) \subseteq_\Box \hat{\nu}(\mathbf{b})$.
10. $x\hat{S}_w y \rightarrow x(\hat{\mathcal{S}}_w \cup \hat{R})^* y$.
    This is clear since $\hat{S} = S$ and $\hat{R} \subseteq R$.
11. $w\hat{K}x\hat{K}^1 y(\hat{\mathcal{S}}_w)^{\mathrm{tr}} y' K^1 z \rightarrow x(\hat{K}^1)^{\mathrm{tr}} z$.
    See the proof that $\hat{F} \models \mathcal{J}_{K^1}$ below.
12. $x\hat{R}y \rightarrow x(\hat{R}^1)^{\mathrm{tr}} y$.
    Immediate.

$\dashv$

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

Lemma 11.23. *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 11.7. 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. $\dashv$

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

Lemma 11.24. *For all x. If $x\hat{K}^1\mathbf{b}$ then $x \subset \mathbf{a}$.*

Proof. Let
$$K' = K \cup \{(x, \mathbf{b}) \mid x\hat{K}\mathbf{b} \wedge x \subset \mathbf{a}\}.$$

We show that K' satisfies the conditions of T in Lemma 11.7. Clearly $K' \subseteq \hat{K}$. Suppose $wK'xK'^1 y(\mathcal{S}_w)^{\mathrm{tr}} y' K'^1 z$. We have to show $xK'z$. Since $K \subseteq K'$ we can assume that $z = \mathbf{b}$ and $wKxK^1 y(\mathcal{S}_w)^{\mathrm{tr}} y'$. Now, by definition of K', $y' \subset \mathbf{a}$. And by definition of $\subset$, $x \subset y'$. So, $x \subset \mathbf{a}$. And thus, since $x\hat{K}\mathbf{b}$ is clear, we conclude $xK'\mathbf{b}$. $\dashv$

Lemma 11.25. *Suppose the conditions of Lemma 11.23 are satisfied and let* $\mathbf{u}$ *be the* u *asserted (by that lemma) to exist. Then for all* $x \neq \mathbf{a}$*,*

$$\textit{If } x\hat{K}^1\mathbf{b} \textit{ then } xK^1\mathbf{u}.$$

Proof. By Lemma 11.24 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 11.17, $xKx_{i+1}K\mathbf{b}$. 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 $i \geq 0$, $x_{n-(i+1)}K^1\mathbf{u}$.

$\dashv$

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

Proof. We only show that $\hat{F}$ satisfies $\mathcal{J}_{K^1}$, $\mathcal{J}_{\subset}$ and $\mathcal{J}_{\nu_1}$. The other invariants are immediate. Let $K = K(\hat{F})$.

$\mathcal{J}_{K^1}$ $w\hat{K}x\hat{K}^1y(\hat{\mathcal{S}}_w)^{\mathrm{tr}}y'\hat{K}^1z \rightarrow x\hat{K}^1z$

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 11.23 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 11.25, $t = \mathbf{a}$ 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 11.17 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 11.25 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.

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

Suppose $x \neq y$, $\hat{F} \models x \subset y$. We can assume that $y = \mathbf{b}$. But $\mathbf{b} \subset x$ is impossible since this would imply that for some t, $\mathbf{b}\hat{K}t$.

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

Suppose '$\hat{\nu}(w, y)$ is defined' and $v\hat{K}y$. We can assume that $y = \mathbf{b}$. We have that $\hat{\nu}(w, \mathbf{b})$ implies $w = \mathbf{a}$. And by Lemma 11.24 we have $v \subset \mathbf{a}$.

$\dashv$

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

Proof. Let $K = K(\hat{F})$. We run through the main-invariants.

$\mathcal{I}_\Box$ 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 11.23 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 11.13, $x\hat{K}^1\mathbf{b} \Rightarrow x \subseteq_\Box \mathbf{a}$.

$\mathcal{I}_{\mathrm{d}}$ $w\hat{K}^1 x \wedge w\hat{K}^{\geq 2}x'(\hat{S}_w \cup \hat{K})^* x \rightarrow$ 'there does not exists a deficiency in w w.r.t. x'
We only have to consider the case $x = \mathbf{b}$. Suppose

$$w\hat{K}^1\mathbf{b}, \tag{25}$$

$$\exists x' \in \hat{F}\,(w\hat{K}^{\geq 2}x'(\hat{S}_w \cup \hat{K})^*\mathbf{b}). \tag{26}$$

(26) implies $\exists x'' \in \hat{F}\ x'(\hat{S}_w \cup \hat{K})^* x''\hat{K}\mathbf{b}$. But then $w\hat{K}^{\geq 2}\mathbf{b}$, in contradiction with (25).

$\mathcal{I}_S$ $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* y' \rightarrow$ 'the $\subseteq_\Box$–max of $\{\nu(t) \mid w\hat{K}t\hat{K}^1 y'\}$, if it exists, is $\subseteq_\Box$–larger than $\nu(x)$'
Suppose $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* y'$. We can assume that $y' = \mathbf{b}$ and thus for some $y'' \in F$, $y(\hat{S}_w \cup \hat{K})^* y''\hat{K}\mathbf{b}$. Since F is an $\mathbf{IL}\mathsf{M}_0$–frame this implies $y\hat{S}_w y''$ and thus $x\hat{K}\mathbf{b}$.

$\mathcal{I}_\mathcal{N}$ $w\hat{K}x\hat{K}y \wedge \hat{F} \models y \in \mathcal{N}_w^A \rightarrow \hat{F} \models x \in \mathcal{N}_w^A$
Suppose $w\hat{K}x\hat{K}y$ and $\hat{F} \models y \in \mathcal{N}_w^A$. We only have to consider the case $y = \mathbf{b}$. Then, by Lemma 11.20, $\hat{\nu}(w, \mathbf{b}) = A$ or for some $t \in \mathcal{N}_w^A$ 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

$$t\hat{K}^1\mathbf{b}. \tag{27}$$

We consider the following cases.

(a) The conditions of Lemma 11.23 are *not* fulfilled.
In this case $t = \mathbf{a}$. Moreover $x\hat{K}^*$. So, $\mathbf{a} \in \mathcal{N}_w^A$ and since $F \models \mathcal{I}_\mathcal{N}$ we conclude $x \in \mathcal{N}_w^A$.

(b) The conditions of Lemma 11.23 are fulfilled.
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 11.22 and $\mathcal{J}_{\mathcal{N}_1}$ we have

$$\mathbf{u} \in \mathcal{N}_w^A \Leftrightarrow \mathbf{a} \in \mathcal{N}_w^A.$$

Also, by (27),

$$\mathbf{u} \in \mathcal{N}_w^A \vee \mathbf{a} \in \mathcal{N}_w^A.$$

So since $x\hat{K}\mathbf{u}$ or $x = \mathbf{a}$ or $x\hat{K}\mathbf{a}$ we obtain $x \in \mathcal{N}_w^A$ by $F \models \mathcal{I}_\mathcal{N}$.

$\mathcal{I}_{\mathcal{D}}$ $x\hat{R}y \to \exists A \in (\hat{\nu}(y) \setminus \hat{\nu}(x)) \cap \{\Box D \mid D \in \mathcal{D}\}$
We have $A \in \mathcal{D}$ and $\Diamond A \in \hat{\nu}(\mathbf{a})$. We have chosen $\Delta_{\mathbf{b}}$ such that $\Box\neg A \in \Delta_{\mathbf{b}}$.

$\dashv$

To finish this subsection we note that by Lemma 11.12 we can extend $\hat{F}$ to an adequate $\mathbf{IL}\mathsf{M}_0$–frame that satisfies all invariants. Moreover, by Corollary 11.19, we know that this frame satisfies all the invariants.

11.6 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. There are two cases to consider.

Case 1: $\mathbf{a}R^1\mathbf{b}$. Pick some $y \notin W$. Pick some formula 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 $A = \bot$.) Let Δ_y be a MCS as given by Lemma 4.21. In particular, $\nu(\mathbf{a}) \prec_A \nu(\Delta_y)$. (By Lemma 2.8, $\prec_\bot = \prec$.) Now we put

$$\begin{aligned}\hat{F} = \langle & W \cup \{y\},\\ & R \cup \{(\mathbf{a}, y)\},\\ & S \cup \{(\mathbf{a}, \mathbf{b}, y)\},\\ & \nu \cup \{(y, \Delta_y), ((\mathbf{a}, y), A)\}\rangle.\end{aligned}$$

Case 2: $\mathbf{a}R^{\geq 2}\mathbf{b}$. 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 $\nu(\mathbf{a}) \prec_A \nu(x)$. (By Lemma 2.8, $\prec_\bot = \prec$.) So we have

$$C \rhd D \in \nu(\mathbf{a}) \prec_A \nu(x) \ni \Diamond C.$$

Apply Lemma 11.15 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

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

In both of the cases above there exists some set Y, disjoint from W, and some set $\{\Delta_y \mid y \in Y\}$ of MCS's such that

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

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

Proof. We run through properties (1.–5.) of Definition 5.1 (quasi–frames) and properties (6.–10.) of Definition 11.9 (quasi–$\mathbf{IL}\mathsf{M}_0$–frames)

1. $\hat{R}$ is conversely well-founded
   This is clear.
2. $x\hat{S}_w y \to w\hat{R}x, y$
   This is clear.
3. $w\hat{R}x \to \hat{\nu}(w) \prec \hat{\nu}(x)$
   Clear by choice of $\hat{\nu}(y)$ for $y \in Y$.
4. $A \neq B \Rightarrow \mathcal{G}_w^A \cap \mathcal{G}_w^B = \emptyset$.
   Suppose $\hat{F} \models y \in \mathcal{G}_w^A \cap \mathcal{G}_w^B$. We only have to consider the case $y \in Y$. If $w = \mathbf{a}$ then $F \models \mathbf{b} \in \mathcal{G}_w^A \cap \mathcal{G}_w^B$. So assume $w \neq \mathbf{a}$. Then for some x with $\hat{\nu}(w, x) = A$ and for some x' with $\hat{\nu}(w, x') = B$ we have the following
   $$x(\hat{R} \cup \hat{S})^* y$$
   $$x'(\hat{R} \cup \hat{S})^* y$$
   But then also
   $$x(\hat{R} \cup \hat{S})^* \mathbf{b}$$
   $$x'(\hat{R} \cup \hat{S})^* \mathbf{b}$$
   And thus
   $$\hat{F} \models \mathbf{b} \in \mathcal{G}_w^A \cap \mathcal{G}_w^B.$$
   Conclusion: $A = B$.
5. $\hat{F} \models x \in \mathcal{C}_w^A \Rightarrow \hat{\nu}(w) \prec_A \hat{\nu}(x)$
   We only have to consider the case $\hat{F} \models y \in \mathcal{C}_w^A$ for $y \in Y$. If $w \neq \mathbf{a}$ then clearly $F \models \mathbf{a} \in \mathcal{C}_w^A$. So, $\hat{\nu}(w) \prec_A \hat{\nu}(\mathbf{a})$ and thus $\hat{\nu}(w) \prec_A \hat{\nu}(y)$.
6. $\hat{K}$ is conversely well-founded.
   Any $\hat{K}$–chain that is not a K–chain must include a world $y \in Y$. But since for such y there is no z for which $y\hat{K}z$, such a chain must be finite.
7. $x\hat{K}y \to \hat{\nu}(x) \prec \hat{\nu}(y)$.
   We can assume $y \in Y$. By Lemma 11.13 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 $\hat{\nu}(x) \prec \hat{\nu}(y)$.
8. $\hat{F} \models x \in \mathcal{N}_w^A \Rightarrow \hat{\nu}(w) \prec_A \hat{\nu}(x)$.
   We only need to consider the cases $\hat{F} \models y \in \mathcal{N}_w^A$ for $y \in Y$. There are two possibilities. For some z, $\hat{F} \models z \in \mathcal{N}_w^A$ and $z\hat{K}y$, or $w = \mathbf{a}$ and $\hat{F} \models \mathbf{b} \in \mathcal{N}_w^A$. Since otherwise $\mathcal{N}_w^A - \{y\}$ satisfies the two conditions of Definition 11.8 but is strictly smaller than $\mathcal{N}_w^A$. In the first case we have $F \models z \in \mathcal{N}_w^A$ and thus $\hat{\nu}(w) \prec_A \hat{\nu}(z)$. By item 7. we obtain $\hat{\nu}(w) \prec_A \hat{\nu}(y)$. in the second case we are done by choice of $\hat{\nu}(y)$.

9. $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* y' \to \hat{\nu}(x) \subseteq_\Box \hat{\nu}(y')$.
   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.
   (a) For some t, $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* t\hat{K}y'$.
       $\hat{\nu}(y')$ is $\subseteq_\Box$-larger than $\hat{\nu}(t)$ by Item 7. above. Also we have $wKxKy(S_w \cup K)^* t$. So, $\hat{\nu}(x) = \nu(x) \subseteq_\Box \nu(t) = \hat{\nu}(t)$.
   (b) $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* \mathbf{b}\hat{S}_w y'$.
       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 $\nu(x)$.
10. $x\hat{S}_w y \to x(\hat{S}_w \cup \hat{R})^* y$.
    Immediate.
11. $w\hat{K}x\hat{K}^1 y(\hat{\mathcal{S}}_w)^{\mathrm{tr}} y' K^1 z \to x(\hat{K}^1)^{\mathrm{tr}} z$.
    See the proof that $\hat{F} \models \mathcal{J}_{K^1}$ below.
12. $x\hat{R}y \to x(\hat{R}^1)^{\mathrm{tr}} y$.
    Immediate.

⊣

Lemma 11.26. *For any $x \in \hat{F}$ and $y \in Y$ we have*

$$x\hat{K}^1 y \to x \subset \mathbf{a}.$$

Proof. Put

$$K' = K \cup \{(x, y) \mid y \in Y,\ x\hat{K}y,\ x \subset \mathbf{a}\}.$$

We will show that $x\hat{K}^1 y \to xK'y$. To this end we show that K' satisfies the conditions of T in Lemma 11.7.

1. $\hat{R} \subseteq K' \subseteq \hat{K}$
   $K' \subseteq \hat{K}$ is clear. Suppose $x\hat{R}y$. We can assume that $y \in Y$. Then $x = \mathbf{a}$ so certainly $x \subset \mathbf{a}$.
2. $wK'xK'^1 y(\hat{\mathcal{S}}_w)^{\mathrm{tr}} y' K'^1 z \to xK'z$.
   Suppose $wK'xK'^1 y(\hat{\mathcal{S}}_w)^{\mathrm{tr}} y' K'^1 z$. Since $K \subseteq K'$ we can assume that at least one of w, x, y, y', z is in Y. The only candidate for this is z. Now, by definition of K', $y'K'^1 z$ implies $y' \subset \mathbf{a}$. By definition of $\subset$, $x \subset y'$ and thus $x \subset \mathbf{a}$.

So if $x\hat{K}^1 y$ then $xK'y$. But if $y \in Y$ then xKy does not hold. So in these cases we must have $x \subset \mathbf{a}$.

⊣

Lemma 11.27. *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^1y. By Lemma 11.26 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 $i \leq n$, $x_{n-i}K^1 z$. $\dashv$

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

Proof. We run trough all the sub-invariants.

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

Suppose $w\hat{K}^{\geq 2}x(\hat{\mathcal{S}}_w)^{\mathrm{tr}}y$ and $w\hat{K}^{\geq 2}x'(\hat{\mathcal{S}}_w)^{\mathrm{tr}}y$. 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 $w\hat{K}^1x$. 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 $wK^{\geq 2}x'(\mathcal{S}_w)^{\mathrm{tr}}\mathbf{b}$. 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 $wK^{\geq 2}x(\mathcal{S}_w)^{\mathrm{tr}}\mathbf{b}$ and $wK^{\geq 2}x'(\mathcal{S}_w)^{\mathrm{tr}}\mathbf{b}$. So, by $F \models \mathcal{J}_{\mathrm{u}}$, we obtain $x = x'$.

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

Suppose $w\hat{K}x\hat{K}^1y(\hat{\mathcal{S}}_w)^{\mathrm{tr}}y'\hat{K}^1z$. 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 $\mathbf{a}K^1a_1$. By Lemma 11.27 we have $y'K^1a_1$ and thus, since $F \models \mathcal{J}_{K^1}$, xK^1a_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.

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

Suppose $x \neq y$ and $\hat{F} \models x \subset y$. We can assume that at least on of x, y is in Y. Since $x \subset y$ implies that for some t, $x\hat{K}t$, this can only be y. But then, for the very same reason, we cannot have $y \subset x$.

$\mathcal{J}_{\mathcal{N}_1}$ $x(\hat{\mathcal{S}}_v)^{\mathrm{tr}}y \wedge w\hat{K}y,\ \hat{F} \models x \in \mathcal{N}^A_w \Rightarrow \hat{F} \models y \in \mathcal{N}^A_w$

Suppose $x(\hat{\mathcal{S}}_v)^{\mathrm{tr}}y$, $w\hat{K}y$ and $\hat{F} \models x \in \mathcal{N}^A_w$. We can assume $v \neq w$ and $y \in Y$. Then $v = \mathbf{a}$ and we can assume that $x = \mathbf{b}$. Moreover, since $\hat{F} \models \mathcal{J}_{\nu_1}$,

$$\hat{F} \models w \subset \mathbf{a}. \tag{28}$$

We will show that $w\hat{K}^{\geq 2}y$. Suppose for a contradiction that $w\hat{K}^1y$. W.l.o.g. we can assume $\mathbf{a}\hat{K}^1\mathbf{b}$. There can be two reasons for $\mathbf{b} \in \mathcal{N}^A_w$.

(a) For some $t \in \mathcal{N}_w^A$, $t\mathcal{S}_w\mathbf{b}$.
This would implie that $wK^1\mathbf{b}$ and thus $\mathbf{a} \subset w$. A contradiction.

(b) For some $t \in \mathcal{N}_w^A$, $tK\mathbf{b}$.
Since $t \in \mathcal{N}_w^A$ we have wKt and thus $wK^{\geq 2}\mathbf{b}$. A contradiction.

So $w\hat{K}^1 y$. Fix some a_1 for which $\mathbf{a}K^1 a_1 K^* \mathbf{b}$. By Lemma 11.27 we have $w\hat{K}^1 a_1$. By $F \models \mathcal{I}_\mathcal{N}$, $a_1 \in \mathcal{N}_w^A$. Now, by Lemma 11.27, $\nu(w, a_1) = A$. (The possibility that for some $t \in \mathcal{N}_w^A$, tKa_1 is excluded since this implies $wK^{\geq 2}a_1$.) This gives, by $F \models \mathcal{J}_{\nu_1}$, $\mathbf{a} \subset w$. In contradiction with (28)

$\mathcal{J}_{\mathcal{N}_2}$ $x(\hat{\mathcal{S}}_w)^{\mathrm{tr}}y$, $\hat{F} \models y \in \mathcal{N}_w^A \Rightarrow \hat{F} \models x \in \mathcal{N}_w^A$

Suppose $x(\hat{\mathcal{S}}_w)^{\mathrm{tr}}$ $\hat{F} \models y \in \mathcal{N}_w^A$. We can assume $y \in Y$. So $w = \mathbf{a}$. The only reason for $\hat{F} \models y \in \mathcal{N}_\mathbf{a}^A$ is $F \models \mathbf{b} \in \mathcal{N}_\mathbf{a}^A$. This implies, by $F \models \mathcal{I}_\mathcal{N}$, $F \models x \in \mathcal{N}_\mathbf{a}^A$ and thus $\hat{F} \models x \in \mathcal{N}_\mathbf{a}^A$.

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

By $\mathcal{J}_{\nu_4}$ and Lemma 11.26.

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

Suppose $\hat{\nu}(w, y)$ is defined. We can assume that $y \in Y$. Then $w = \mathbf{a}$ and thus $w\hat{K}^1 y$.

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

Suppose both $\hat{\nu}(v, y)$ and $\hat{\nu}(w, y)$ are defined. We can assume that $y \in Y$. But then $v = \mathbf{a} = w$.

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

Suppose $x(\hat{\mathcal{S}}_w)^{\mathrm{tr}}y$. We can assume that $y \in Y$. So $w = \mathbf{a}$, and $\hat{\nu}(w, y)$ is defined.

$\dashv$

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

Proof. We run through the main-invariants.

$\mathcal{I}_\Box$ For all y, $\{\hat{\nu}(x) \mid x\hat{K}^1 y\}$ is linearly ordered by $\subseteq_\Box$.

Let $y \in \hat{F}$ and consider the set $\{x \mid xK^1 y\}$. Since $\hat{K} \upharpoonright_F = K$ and for all $y \in Y$ there does not exists z with $y\hat{K}^1 z$ we only have to consider the case $y \in Y$. Fix some a_1 such that $\mathbf{a}K^1 a_1 K^* \mathbf{b}$. By Lemma 11.26 for any such y we have

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

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

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

Suppose $w\hat{K}^{\geq 2}y'\hat{S}_w y$, $w\hat{K}y$ and $w\hat{K}^1 y$. If $y \in Y$ then $w = \mathbf{a}$ and we have solved the deficiency in $\mathbf{a}$ w.r.t. $\mathbf{b}$ using case 2 (as described on page 85). But then by construction there are no deficiencies in w w.r.t. y. So suppose $y \in F$. Then also $w, x, x', y' \in F$ and in F there does not exists a deficiency in w w.r.t. y. The only reason for there to exist one in $\hat{F}$ is that wKy does not hold. Which is not so.

$\mathcal{I}_S$ $w\hat{K}x\hat{K}y(\hat{S}_w \cup \hat{K})^* y' \rightarrow$ 'the $\subseteq_\Box$–max of $\{\nu(t) \mid w\hat{K}t\hat{K}^1 y'\}$, if it exists, is $\subseteq_\Box$–larger than $\nu(x)$'

Suppose $w\hat{K}x\hat{K}y(\hat{S}_w)^{\mathrm{tr}} y'$. We can assume that $y' \in Y$. Thus $w = \mathbf{a}$. So, $w\hat{K}^1 y'$. In other words, $\{t \mid wKtKy'\} = \emptyset$.

$\mathcal{I}_\mathcal{N}$ $w\hat{K}x\hat{K}y \wedge \hat{F} \models y \in \mathcal{N}^A_w \rightarrow \hat{F} \models x \in \mathcal{N}^A_w$

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 11.26, $x \subset \mathbf{a}$. So, $wKxK\mathbf{b}$. By Lemma 11.22, $F \models \mathbf{b} \in \mathcal{N}^A_w$ and thus $\hat{F} \models x \in \mathcal{N}^A_w$.

$\mathcal{I}_\mathcal{D}$ $x\hat{R}y \rightarrow \exists A \in (\hat{\nu}(y) \setminus \hat{\nu}(x)) \cap \{\Box D \mid D \in \mathcal{D}\}$

In the case $\mathbf{a}R^1\mathbf{b}$ we have chosen $\hat{\nu}(y)$ to contain $\Box\neg C$. In the case $\mathbf{a}R^{\geq 2}\mathbf{b}$ we have for each $y \in Y$, that $\hat{\nu}(t) \subseteq_\Box \hat{\nu}(y)$ for some (or actualy each) $\mathbf{a}RtR\mathbf{b}$.

$\dashv$

To finish this section we noting that by Lemma 11.12 we can extend $\hat{F}$ to an adequate $\mathbf{IL}\mathsf{M}_0$–frame that satisfies all invariants. Moreover, by Corollary 11.19, we know that this frame satisfies all the invariants.

11.7 Rounding up

We have to show that the union of a bounded chain of frames that satisfy all the invariants is an $\mathbf{IL}\mathsf{M}_0$–frame. But the $\mathbf{IL}\mathsf{M}_0$–frame conditions are part of the invariants and 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.

11.8 Considerations

In this section we discus some ideas which might simplify the completeness proof for $\mathbf{IL}\mathsf{M}_0$.

During the construction of the $\mathbf{IL}\mathsf{M}_0$–closure, the set $\mathcal{C}^A_w$ is not a constant object. Therefore showing that $\mathcal{C}^A_w$ lies A–critical above w is not that trivial as in the case for $\mathbf{IL}$. For this reason, the set $\mathcal{N}^A_w$ and the relation K was introduced. The whole exposition might be easier if we define $\mathcal{C}^A_w$ as follows.

1. $w \prec_A x$ and $x \in \mathcal{G}^A_w$, imply $x \in \mathcal{C}^A_w$
2. $\mathcal{C}^A_w$ is closed under R and S_w

However, K is also used in ensuring the conversely well–foundedness of R. So, this does not yet completely eliminate the need for this object yet. Tough at finite and transitive frame this is not really an issue since there conversely well–foundedness equals irreflexiveness.

The use of the relations $\mathcal{S}_w$ can be circumvented using the following observations.

1. S_w relations that are not $\mathcal{S}_w$ only start to emerge when closing of under the property $wRxRy \rightarrow xS_wy$.

2. If a frame does not satisfy $wRxRy \to xS_wy$, but does satisfy all the other $\mathbf{IL}\mathsf{M}_0$–frame conditions. Then we can close it of, such that no (new) problems nor deficiencies occur. (One basically has to show that we do not need any new R relations when we close off.)

So, if we perform the construction method while ignoring the condition $wRxRy \to xS_wy$. Then all S_w relations are $\mathcal{S}_w$ (so there is no need to call them as such). And at the end we can extend our problem–less and deficiency–less frame to a problem–less and deficiency–less $\mathbf{IL}\mathsf{M}_0$–frame.

In the completeness proof for $\mathrm{IL}\mathsf{M}_0$ we used a complicated existence lemma that made us add, in a single step of the construction, a whole bunch of worlds to our model. Additionally we need to keep an invariant of the form 'If $wRtRyS_wy'$ then there are no deficiencies in w w.r.t. y''. Here we 'localize' that complicated existence lemma. The result of this, is that it suffices to add only one world in each step and we can forget about the above invariant.

MCS means $\mathrm{IL}\mathsf{M}_0$ maximal consistent and $\vdash$ is provability in $\mathrm{IL}\mathsf{M}_0$.

Definition 11.28 ($\prec_S$). Let S be a set of formulas. We write $\Gamma \prec_S \Delta$ iff for all formulas C and $S_0, \ldots, S_{n-1} \in S$, $C \rhd \bigvee_{i<n} \neg S_i \in \Gamma$ implies $\neg C, \Box\neg C \in \Delta$.

In this context we have $\Gamma \prec_A \Delta$ iff. $\Gamma \prec_S \Delta$ for some S with $\neg A \in S$. Also notice that if $S' \subseteq S$ and $\Gamma \prec_S \Delta$ then $\Gamma \prec_{S'} \Delta$.

Lemma 11.29. *Suppose $\Gamma \prec_S \Delta$, $C \rhd D \in \Gamma$, $C \in \Delta$. There exists a Δ' with $\Gamma \prec_S \Delta'$ and $B \in \Delta'$.*

Proof. Let

$$\mathcal{D} = \{C \mid \exists S_0, \ldots, S_{n-1} \in S\ (C \rhd \bigvee_{i<n} \neg S_i \in \Gamma)\}.$$

Suppose that $\{B\} \cup \{\neg C, \Box\neg C \mid C \in \mathcal{D}\}$ is inconsistent. Then for some finite $\mathcal{D}' \subseteq \mathcal{D}$ we have

$$\vdash B \to \bigvee_{C \in \mathcal{D}'} C \vee \Diamond C.$$

So,

$$\vdash B \rhd \bigvee_{C \in \mathcal{D}'} C.$$

But then for some $S_0, \ldots S_{k-1} \in S$ we have $B \rhd \bigvee_{i<k} S_k \in \Gamma$. So, since $A \rhd B \in \Gamma$ we obtain

$$A \rhd \bigvee_{i<k} S_k \in \Gamma.$$

So, since $\Gamma \prec_S \Delta$, we conclude $\neg A, \Box\neg A \in \Delta$. A contradiction. $\dashv$

Lemma 11.30. *Suppose $\Gamma \prec_S \Delta$, $A \rhd B \in \Gamma$ and $\Diamond A \in \Delta$. There exists Δ' with $\Gamma \prec_{S \cup \{\Box E \mid \Box E \in \Delta\}} \Delta'$ and $B \in \Delta'$.*

Proof. Let

$$\mathcal{D} = \{C \mid \exists S_0, \ldots, S_{n-1} \in S \cup \{\Box E \mid \Box E \in \Delta\}\ (C \rhd \bigvee_{i<n} S_i \in \Gamma)\}.$$

Suppose $\{B\} \cup \{\neg C, \Box\neg C \mid C \in \mathcal{D}\}$ is inconsistent. Then for some finite $\mathcal{D}' \subseteq \mathcal{D}$ we have

$$\vdash B \to \bigvee_{C \in \mathcal{D}'} C \vee \Diamond C.$$

So,

$$\vdash B \rhd \bigvee_{C \in \mathcal{D}'} C.$$

But then for some $S_0, \ldots, S_k \in S$ and $\Box E_0, \ldots, \Box E_l \in \Delta$ we have $B \rhd \bigvee_{i<k} S_i \vee \bigvee_{j<l} \Box E_l \in \Gamma$. So, since $A \rhd B \in \Gamma$ and by an application of the M_0 axiom we obtain

$$\Diamond A \wedge \bigwedge_{j<l} \Box E_j \rhd (\bigvee_{i<k} S_i \vee \bigvee_{j<l} \Box E_l) \wedge \bigwedge_{j<l} \Box E_j \in \Gamma.$$

Since $\Gamma \prec_S \Delta$ this implies $\Diamond A \notin \Delta$ or $\bigwedge_{j<l} \Box E_j \notin \Delta$. A contradiction. $\dashv$

We should maintain (variantions of) the following invariants (where we identify worlds t with their labels $\nu(t)$).

1. $wRxRyS_wy' \to w \prec_{\{\Box D \mid \Box D \in x\}} y'$
2. $yS_wy' \wedge w \prec_S y \to w \prec_S y'$.

If $x \prec_S y'$ then $S \subseteq y'$ So, invariant 1. ensures that $wRxRyS_wy' \to x \subseteq_\Box y'$. Lemma 11.30 ensures that we can solve deficiencies while preserving both the Invariants 1. and 2.. Invariant 2. ensures that invariant 1. is preserved when we take S_w transitive. Lemma 11.29 ensures that we can solve deficiencies while preserving Invariant 2..

12 The logic $\mathbf{IL}\mathsf{W}^*$

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

Theorem 12.1. $\mathbf{IL}\mathsf{W}^*$ *is a complete logic.*

Instead of applying the construction method in full we indicate how we can modify the completeness proof of $\mathbf{IL}\mathsf{M}_0$.

12.1 Preliminaries

Definition 12.2 ($\subsetneq_\Box^{\mathcal{D}}$)**.** 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 12.3. *Let F be an quasi–frame and $\mathcal{D}$ be a finite set of formulas. If $wRxRyS_wy' \to \nu(x) \subsetneq^{\mathcal{D}}_{\Box} \nu(y')$ then $(R;S_w)$ is conversely well–founded.*

Proof. Take $w \in F$ and let $x_0(R;S_w)x_1(R;S_w)x_2\cdots$ be an $R;S_w$ chain in F. For all $i \geq 1$ we have wRx_i. Thus, for all $i \geq 1$, $x_i \subsetneq^{\mathcal{D}}_{\Box} x_{i+1}$. Since $\mathcal{D}$ is finite the chain must be finite as well. ⊣

Lemma 12.4. *Let F be a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. If $wRxRyS_wy' \to \nu(x) \subsetneq^{\mathcal{D}}_{\Box} \nu(y')$ then $wRxRy(S_w \cup R)^*y' \to \nu(x) \subsetneq^{\mathcal{D}}_{\Box} \nu(y')$*

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

Definition 12.5 (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 hold.

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

Definition 12.6 (Quasi–$\mathbf{IL}\mathsf{W}^*$–frame). Let $\mathcal{D}$ be a set pf formulas. We say that a quasi–$\mathbf{IL}\mathsf{M}_0$–frame is a quasi–$\mathbf{IL}\mathsf{W}^*$–frame (w.r.t. $\mathcal{D}$) iff. the following additional property hold.

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

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

Corollary 12.7. *For any adequate $\mathbf{IL}\mathsf{W}^*$–frame F and for each $w \in F$ we have that $(R;S_w)$ is conversely well–founded*

Lemma 12.8. *Any quasi–$\mathbf{IL}\mathsf{W}^*$–frame can be extended to an adequate $\mathbf{IL}\mathsf{W}^*$–frame. (Both w.r.t. the same set of formulas $\mathcal{D}$.)*

Proof. Let F be a quasi–$\mathbf{IL}\mathsf{W}^*$–frame. Then in particular F is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. So consider the proof of Lemma 11.12. There we constructed a sequence of quasi–$\mathbf{IL}\mathsf{M}_0$–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–$\mathbf{IL}\mathsf{W}^*$–frame then each F_i is as well. And additionally that $\hat{F}$ is an adequate $\mathbf{IL}\mathsf{M}_0$–frame.

But this is rather trivial. As noted in the proof of Lemma 11.12, 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–$\mathbf{IL}\mathsf{W}^*$–frame.

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

⊣

We finish the preliminaries with an adaption of Lemma 11.15 to the logic $\mathbf{IL}\mathsf{W}^*$.

Lemma 12.9. *Let Γ and Δ be* MCS*'s. $\Gamma \prec_C \Delta$.*

$$P \rhd Q, S_1 \rhd T_1, \ldots, S_n \rhd T_n \in \Gamma$$

and

$$\Diamond P \in \Delta.$$

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. 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

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

and, for all $i \notin I$,

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

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

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

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

$$\Box\neg P, Q, \overline{S}_I, \Box A, \neg B \wedge \Box\neg B \vdash \bot \tag{29}$$

or,

$$\text{for some } i \notin I, \quad \Box\neg P, T_i, \overline{S}_I, \Box A, \neg B \wedge \Box\neg B \vdash \bot. \tag{30}$$

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

$$\Box\neg P, T_{i_k}, \overline{S}_{I_k}, \Box A, \neg B \wedge \Box\neg B \vdash \bot. \tag{31}$$

Additionally, we will verify that for each k

(29) does not hold with I_k for I.

We will define i_k with induction on k. We define $I_1 = \emptyset$. And by Lemma 4.21, (29) does not hold with $I = \emptyset$. Moreover, because of this, (30) must be true with $I = \emptyset$. So there exist some $i \in \{1, \ldots, n\}$ such that

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

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

$$\Box\neg P, Q, \overline{S}_{I_{k+1}}, \Box A, \neg B \wedge \Box\neg B \nvdash \bot. \tag{32}$$

Let us suppose that this is not so. Then

$$\vdash \Box(Q \to \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_k}). \tag{33}$$

So,

$$\begin{array}{lll}
\Gamma \vdash P \rhd Q & & \\
& \rhd \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee S_{i_k} & \text{by (33)} \\
& \rhd \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee T_{i_k} & \\
& \rhd \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee (T_{i_k} \wedge \Box\neg P \wedge \Box A \wedge \neg B \wedge \Box\neg B \wedge \overline{S}_{I_k}) & \\
& \rhd \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} & \text{by (31)} \\
& \quad \vdots & \\
& \rhd \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} & \\
& \rhd \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B \vee T_{i_1} & \\
& \rhd \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B \vee (T_{i_1} \wedge \Box\neg P \wedge \Box A \wedge \neg B \wedge \Box\neg B) & \\
& \rhd \Diamond P \vee \Diamond\neg A \vee B \vee \Diamond B. & \text{by (31), with } k = 1.
\end{array}$$

So, by W, $P \rhd \Diamond\neg A \vee B \vee \Diamond B$. And thus by 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 4.21 there exists some MCS Δ that contains $B \vee \Diamond B$. But, since $\Gamma \prec_C \Delta$, by that very same Lemma we can assure that $\Gamma \prec_C \Delta'$. And this implies $\neg B \wedge \Box\neg B \in \Delta'$. A contradiction. So we have shown (32).

But now, since (32) is indeed true and thus (29) with I_{k+1} for I is false, there must exist some $i \notin I_{k+1}$ such that

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

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

It is clear that for $I = \{1, 2, \ldots, n\}$, (30) cannot be true. Thus, for $I = \{1, 2, \ldots, n\}$, (29) must be true. This implies

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

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. $\dashv$

12.2 Frame condition

We can define a new principle M_0^* as follows.

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

Lemma 12.10. $\mathbf{IL}\mathsf{M}_0\mathsf{W} = \mathbf{IL}\mathsf{W}^* = \mathbf{IL}\mathsf{M}_0^*$

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

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

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 $\mathbf{IL}\mathsf{M}_0\mathsf{W} \vdash \mathsf{M}_0^*$. For, reason in $\mathbf{IL}\mathsf{M}_0\mathsf{W}$ 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. $B \wedge \Box C \rhd (B \wedge \Box C \wedge \Box\neg A) \vee (\Box C \wedge \Diamond A) \rhd B \wedge \Box C \wedge \Box\neg A$.

⊣

Theorem 12.11. *For any* **IL***-frame* F *we have that*

$$F \models \mathsf{W} \Leftrightarrow \forall w\ (S_w; R) \textit{ is conversely well-founded}$$

Proof. Let F be an **IL**-frame.

($\Leftarrow$) Suppose that for all w, $(S_w; R)$ is conversely well-founded. Let $\overline{F}$ be some model based on F. Pick $w \in \overline{F}$ and suppose $w \Vdash A \rhd B$. Pick $x \in \overline{F}$ such that wRx and $x \Vdash A$. There exists some y with xS_wy and $y \Vdash B$. Assume for a contradiction that for no such y we have $y \Vdash \Box\neg A$. Construct a sequence $x_0, x_1, x_2, \ldots$ for which $x_i(S_w; R)x_{i+1}$ and $x_i \Vdash A$ as follows. $x_0 = x$. Suppose x_n is defined. Pick some y for which x_nS_wy and $y \Vdash B$. By transitivity of S_w and the inclusion of S_w in R we have xS_wy. So, by assumption $y \Vdash \Diamond A$. Pick x_{n+1} such that yRx_{n+1} and $x_{n+1} \Vdash A$. But the existence of such a chain is in contradiction with the conversely well-founded-ness of $(S_w; R)$.

($\Rightarrow$) Suppose $F \models \mathsf{W}$. Let $w \in F$ and suppose, for a contradiction, that $(S_w; R)$ is not conversely well-founded. So, let $x_0(S_w; R)x_1(S_w; R)x_2 \cdots$ be an infinite $(S_w; R)$ sequence. Moreover let us choose $y_0, y_1, \ldots$ such that $x_0S_wy_0Rx_1S_wy_1Rx_2 \cdots$. Define the model $\overline{F}$ based on F as follows

$$v \Vdash p \Leftrightarrow v \in \{x_0, x_1, \ldots\}$$
$$v \Vdash q \Leftrightarrow v \in \{y_0, y_1, \ldots\}$$

Now, by definition of $\Vdash$, $w \Vdash p \rhd q$ and thus, since $F \models \mathsf{W}$, $w \Vdash p \rhd q \wedge \Box\neg p$. But also, by definition of $\Vdash$, $w \not\Vdash p \rhd q \wedge \Box\neg p$. A contradiction.

⊣

Corollary 12.12. *For any* **IL** *frame we have that* $F \models \mathsf{W}^*$ *iff.*

1. *For each* w, $(S_w; R)$ *is conversely well-founded.*
2. *For all* w, x, y, y', z *we have* $wRxRyS_wy'Rz \rightarrow xRz$.

A completeness theorem for $\mathbf{IL}\mathsf{W}^*$, however, does not trivially follow from the completeness theorems for $\mathbf{IL}\mathsf{M}_0$ and $\mathbf{IL}\mathsf{W}$ and Lemma 12.10.

Before we move on let us ponder a little on the following. It is a natural question to ask if there is no frame condition for W^* that is more natural than just the conjunction of those of M_0 and W. We do not know of such a condition though. As the following lemma shows we do have one when we consider finite frames only.[26]

Lemma 12.13. *For finite frames F we have that $F \models \mathsf{W}^*$ iff. $(R;S)$ is conversely well-founded and $uRvRwS_uxRy \rightarrow vRy$.*

12.3 Invariants

Let $\mathcal{D}$ be some finite set of formulas closed under subformulas and single negation. As invariants we take all invariants of $\mathbf{IL}\mathsf{M}_0$ and additionally

$\mathcal{I}_{w^*}$ The conditions for an adequate $\mathbf{IL}\mathsf{W}^*$ (w.r.t. $\mathcal{D}$) hold

12.4 Problems

We have to show that we can solve problems an adequate $\mathbf{IL}\mathsf{W}^*$–frame in such a way that we end up with a quasi–$\mathbf{IL}\mathsf{W}^*$–frame. If we have such a frame then in particular it is an $\mathbf{IL}\mathsf{M}_0$–frame. So, as we have seen we can extend this frame to a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. It is easy to see that whenever we started with an adequate $\mathbf{IL}\mathsf{W}^*$–frame we end up with a quasi $\mathbf{IL}\mathsf{W}^*$–frame.

12.5 Deficiencies

We have to show that we can solve any deficiency in an adequate $\mathbf{IL}\mathsf{W}^*$–frame such that we end up with an quasi–$\mathbf{IL}\mathsf{W}^*$–frame. It is easily seen that the process as described in the case of $\mathbf{IL}\mathsf{M}_0$ works if we use Lemma 12.9 instead of Lemma 11.15.

12.6 Rounding up

We have to show that the union of a bounded chain of frames that satisfy all the invariants is an $\mathbf{IL}\mathsf{W}^*$–frame. All the $\mathbf{IL}\mathsf{M}_0$–frame conditions (which are part of the $\mathbf{IL}\mathsf{W}^*$ conditions) are part of the invariants. 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. So we only 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 $\mathcal{I}_{w^*}$, to be precise by the property $wRxRyS_wy' \rightarrow \nu(x) \subsetneq^{\mathcal{D}}_{\Box} \nu(y)$, in exactly the same manner as we can ensure the conversely well–foundedness of R via the boundedness condition. Namely, in any $(R;S_w)$ chain the number of $\Box$–formulas from $\mathcal{D}$ increases. But $\mathcal{D}$ is finite. So, such a chain must be finite as well.

[26] The finiteness condition can be shown to be necessary.

13 Incompleteness results

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 sufficent to distinghuis 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 Vukovicć [Vuk96], Joosten [Joo98] and Verbrugge and was succesfully used to show independce of certain extensions of **IL**.

In this section we set both the generalized Veltman semantics and the model/bisimulation method to work in order to distinghuish some extensions of **IL**, which are indistinguishable using Veltman frames alone.

13.1 Generalized semantics

We use a slight variantion of the semantics used in [Vuk96] to distinguish $\mathbf{IL}\mathsf{P}_0$ from $\mathbf{IL}\mathsf{M}_0$. Although we could have done this with the original semantics, we also compute the characteristic (generalized) frame class for some new interpretability priniciple R. Something we did not succeed in with the original semantics.

Definition 13.1 ($\mathbf{IL}_{\mathbf{set}}$-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 13.2 ($\mathbf{IL}_{\mathbf{set}}$-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 boolen 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 13.3. *Let F be an $\mathbf{IL}_{set}$-frame, $\mathbf{IL}'$ an extionsion of $\mathbf{IL}$ and suppose that for any axiom X of $\mathbf{IL}'$ we have $F \models X$. Then for all A, $\mathbf{IL}' \Vdash A$ implies $F \models A$.*

Proof. Validity on a frame is preserved under modes-ponens and nessecitation. ⊣

Let V be the class of structures of the form $\langle W, R, S\rangle$. Where W is a set, $R \subseteq W \times W$ and $S \subseteq W \times W \times (\mathcal{P}(W) - \emptyset)$. It somewhat dificult to characterize the subclass $I \subseteq V$ such that $F \in I$ iff. $F \models \mathbf{IL}$. And indeed the class if $\mathbf{IL}_{\mathrm{set}}$-frames does not equal I. (For instance the assumption $F \models \mathsf{J4}$ only forces that xS_wY implies wRx and for *s*ome $y \in Y$ we have wRy.)

According to the following lemma, the $\mathbf{IL}_{\mathrm{set}}$-frames do, however, form a subclass of I.

Lemma 13.4 (Soundness of IL). *If* $\mathbf{IL} \Vdash A$ *then for any* $\mathbf{IL}_{set}$*-frame* F*,* $F \models A$*.*

Proof. By Lemma 13.3 it is enough to show that all axioms of $\mathbf{IL}$ are valid on each $\mathbf{IL}_{\mathrm{set}}$-frame.

Showing that all axioms of $\mathbf{GL}$ are valid on each $\mathbf{IL}_{\mathrm{set}}$ frame goes exactly the same as showing that all axioms of $\mathbf{GL}$ are valid in transitive and conversely well-founded (ordinary) Kripke-frames.

Let $F = \langle W, R, S\rangle$ be an $\mathbf{IL}_{\mathrm{set}}$-frame, $\overline{F} = \langle W, R, S, \Vdash\rangle$ a model based on this frame and $w \in W$.

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

Suppose $w \Vdash \Box(A \to B)$. Pick some x with wRx and $x \Vdash A$. Then $x \Vdash B$. Since $xS_w\{x\}$ there exists some Y (namely $\{x\}$) such that xS_wY and $Y \Vdash B$.

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.

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

Suppose $w \Vdash A \rhd C$ and $w \Vdash B \rhd C$. Pick some x with wRx and $x \Vdash A \vee B$. First assume $x \Vdash A$. Now, for some Y, we have xS_wY and $Y \Vdash C$. The case $x \Vdash B$ goes completely simmilair.

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

Suppose $w \Vdash A \rhd B$. and $w \Vdash \Diamond A$. Then for some x with wRx we have $x \Vdash A$. So there exists some Y, with xS_wY, and $Y \Vdash B$. Moreover for each of those y, wRy. But Y is non-empty. Thus for some y we have wRy and $y \Vdash B$.

J5 $\Diamond A \rhd A$

Pick some x such that wRx and $x \Vdash \Diamond A$. Then for some y with xRy we have $y \Vdash A$. Since we have $wRxRy$ we also have $xS_w\{y\}$. So there exists some Y (namely $\{y\}$) with xS_wY and $Y \Vdash A$.

$\dashv$

For the sake of completeness we show completeness of **IL** w.r.t. $\mathbf{IL}_{\mathrm{set}}$-frames.

Theorem 13.5 (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 **IL**-model $M = \langle W, R, S \rangle$, and some $m \in M$ such that $M, m \Vdash \neg A$. Define the $\mathbf{IL}_{\mathrm{set}}$-model $M' = \langle W', R', S', \Vdash' \rangle$ as follows.

$$\begin{aligned} \Vdash' &= \Vdash \text{ on propositional variables and is extended as usual} \\ W' &= W \\ R' &= R \\ S' &= \{(w, x, Y) \mid \forall y \in Y \; xS_w y\} \end{aligned}$$

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 $xS'_w X$, $y \in X$ and $yS'_w Y$. (We can assume $y \notin Y$, but we won't use this.) Pick $y' \in Y$. Then $xS_w y$ and $yS_w y'$. Thus $xS_w y'$. Since $y' \in Y$ was arbitrary we conclude $xS'_w Y$.

We will now see that for all B and all $w \in W$,

$$w \Vdash' B \Leftrightarrow w \Vdash B. \tag{34}$$

Induction on B. Let $w \in W$. The propositional and boolean cases are trivial. The case $B = \Box C$ is easy as well. So, suppose $B = C \rhd D$.

($\Leftarrow$) Suppose $w \Vdash C \rhd D$. We have to show that $w \Vdash' C \rhd D$. Take some x such that $wR'x$ and $x \Vdash' C$. We have to find some Y such that $xS'_w Y$ and $Y' \Vdash' D$. By (IH) we have $x \Vdash C$. $R = R'$ thus for some y we have $xS_w y$ and $y \Vdash D$. By (IH) we also have $y \Vdash' D$. It is thus sufficient to take $Y = \{y\}$.

($\Rightarrow$) Suppose $w \Vdash' C \rhd D$. We have to show that $w \Vdash C \rhd D$. So, take some x for which wRx and $x \Vdash C$. We have to find some y with $xS_w y$ and $y \Vdash D$. By (IH) we have $x \Vdash' C$. So, since $R = R'$ we have $wR'x$ and thus we can thus find some Y with $xS'_w Y$ and $Y \Vdash D$. It is sufficient to take for y any member of Y.

We have shown (34). Thus we have $m \Vdash' \neg A$ and in particular A is not valid on the underling frame of M'.

$\dashv$

Definition 13.6 ($\mathbf{IL}_{\mathbf{set}}\mathsf{M}_0$-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 $wRxRyS_w Y$ there exists some $Y' \subseteq Y$ such that

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

Lemma 13.7. *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.

$\dashv$

Definition 13.8 ($\mathbf{IL}_{\mathbf{set}}\mathsf{P}_0$-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 exsists some $z \in Z$ with yRz,

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

Lemma 13.9. *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.

$\dashv$

Lemma 13.10. *There exists an $\mathbf{IL}_{set}\mathsf{P}_0$-frame which is not an $\mathbf{IL}_{set}\mathsf{M}_0$-frame.*

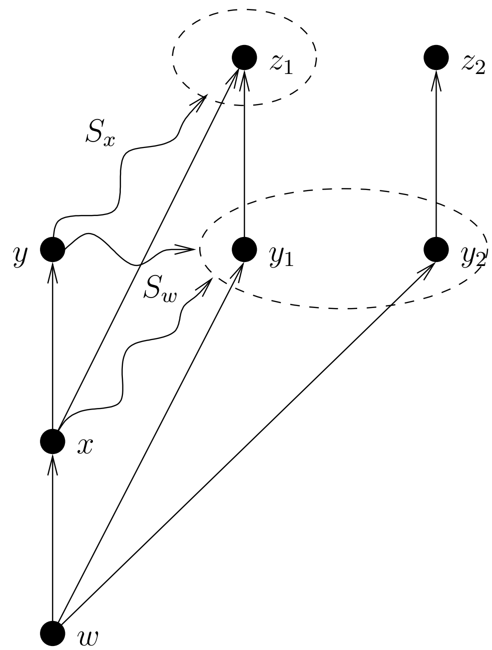

Figure 5: An $\mathbf{IL}_{\text{set}}\mathsf{P}_0$-frame which is not an $\mathbf{IL}_{\text{set}}\mathsf{M}_0$-frame.

Proof. Consider Figure 5. It represents an $\mathbf{IL}_{\text{set}}$-frame. For clarity we have omited 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. Additionaly, 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 $\mathbf{IL}_{\text{set}}\mathsf{P}_0$-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 $\mathbf{IL}_{\text{set}}\mathsf{M}_0$-frame. Put $Y = \{y_1, y_2\}$. We have $wRxRyS_wY$. So, if we do have an $\mathbf{IL}_{\text{set}}\mathsf{M}_0$-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 13.11. $\mathbf{IL}\mathsf{P}_0 \nvdash \mathsf{M}_0$.

Proof. If $\mathbf{IL}\mathsf{P}_0 \vdash \mathsf{M}_0$ then M_0 is valid on any $\mathbf{IL}_{\text{set}}\mathsf{P}_0$-frame. But then any $\mathbf{IL}_{\text{set}}\mathsf{P}_0$-frame is an $\mathbf{IL}_{\text{set}}\mathsf{M}_0$-frame. Which, by Lemma 13.10 is not so. ⊣

In the next section we will investigate a new schema called R.

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

Let us calculate its $\mathbf{IL}_{\text{set}}$-frame class. First a preliminary definition.

Definition 13.12. Let $F = \langle W, R, S\rangle$ be an $\mathbf{IL}_{\text{set}}$-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 13.13. Let $F = \langle W, R, S\rangle$ be an $\mathbf{IL}_{\text{set}}$-frame. We say that F is an $\mathbf{IL}_{\text{set}}\mathsf{R}$-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 13.14. *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 $\mathbf{IL}_{\text{set}}\mathsf{R}$-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 $x \Vdash \neg(A \rhd B) \wedge (\neg C \rhd D)$. We have to find some Y' with xS_wY' and $Y' \Vdash B \wedge \Box C$. 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 $u \Vdash \neg D$. Let Γ be a choice set for (x, y) such that $\Gamma \Vdash \neg D$ 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 al z, $y'Rz$ implies $z \in \Gamma$. So, we are done if we can show that $\Gamma \Vdash C$. So, let $g \in \Gamma$ and suppose $g \Vdash \neg C$. Then xRg and thus since $x \Vdash \neg C \rhd D$, we can find some V such that gS_xV and $V \Vdash D$. But $\Gamma \Vdash \neg D$. So, since $g \in \Gamma$ we have $g \notin V$. By quasi-transitivity we obtain yS_xV. In contradiction with the choice of y.

($\Leftarrow$). 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 $\mathbf{IL}_{\text{set}}$-model $\overline{F} = \langle W, R, S, \Vdash\rangle$ as follows.

$$
\begin{aligned}
v \Vdash p &\Leftrightarrow v = y\\
v \Vdash q &\Leftrightarrow v \in Y\\
v \Vdash t &\Leftrightarrow v \notin \Gamma\\
v \Vdash s &\Leftrightarrow v \in \Gamma
\end{aligned}
$$

Now, $w \Vdash p \rhd q$. So, $w \Vdash \neg(p \rhd t) \wedge (\neg s \rhd t) \rhd q \wedge \Box s$. Also, $\Gamma \Vdash \neg t$. So, $x \Vdash \neg(p \rhd t)$. Moreover $x \Vdash \Box(\neg s \leftrightarrow t)$ so $x \Vdash \neg s \rhd t$. 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 $z \Vdash s$. So, by definition of $\Vdash$, $z \in \Gamma$.

$\dashv$

Before we move on let us consider the definition of an $\mathbf{IL}_{\text{set}}$-frame (Definition 13.1) again. As we have shown, this definition is sufficient for all theorems of $\mathbf{IL}$ to be valid on these frames. There are however other posibilities for the definition of an $\mathbf{IL}_{\text{set}}$-frame. They do not give us all the results we want, however.

For the sake of illustration and because we can then see why we have made the choices we did, let us run through the worst alternative version of the definition of an $\mathbf{IL}_{\text{set}}$-frame.

First we should mention that in [Vuk96] even another version is used. The difference is that definition of quasi-transitivity in these papers runs as follows. xS_wX and for all $y \in X$, yS_wY_y implies $xS_w \bigcup_{y\in X} Y_y$. If we modify the definition of $\mathbf{IL}_{\rm set}$-frame such that it uses this notion of quasi-transitivity then we can also distinguish $\mathbf{IL}\mathsf{M}_0$ and $\mathbf{IL}\mathsf{P}_0$. But we have not succeeded in determining the frame condition for $\mathbf{IL}\mathsf{R}$.

Definition 13.15 ($\mathbf{IL}^*_{\rm set}$-frame). A structure $\langle W, R, S\rangle$ is an $\mathbf{IL}^*_{\rm set}$-frame iff. it is an $\mathbf{IL}_{\rm set}$-frame, *except* that we replace 3.(c) with

3.(c)' If xS_wY, then for some $y \in Y$ we have that for all Z, if yS_wZ then xS_wZ.

$\mathbf{IL}^*_{\rm set}\mathsf{P}_0$-frames and $\mathbf{IL}^*_{\rm set}\mathsf{M}_0$-frames are defined similair to $\mathbf{IL}_{\rm set}\mathsf{P}_0$-frames and $\mathbf{IL}_{\rm set}\mathsf{M}_0$-frames. Except that we take as basis the $\mathbf{IL}^*_{\rm set}$-frames instead of the $\mathbf{IL}_{\rm set}$-frames. The notion of an $\mathbf{IL}^*_{\rm set}$-model is also defined similair as the notion of an $\mathbf{IL}_{\rm set}$-model.

The following lemma is crucial in showing that the $\mathbf{IL}^*_{\rm set}$-frames does not give us much more then the ordinary Veltman frames. It is also easy to see that this lemma is not valid for $\mathbf{IL}_{\rm set}$-frames.

Lemma 13.16. *Let $F = \langle W, R, S\rangle$ be an $\mathbf{IL}^*_{set}$-frame. If xS_wY then for some $y \in Y$ we have $xS_w\{y\}$.*

Proof. Suppose xS_wY. For all $y \in Y$ we have wRy and thus $yS_w\{y\}$. Thus, by 3.(c)', for some $y \in Y$ we must have $xS_w\{y\}$. ⊣

Theorem 13.17. *Let $M = \langle W, R, S, \Vdash\rangle$ be an $\mathbf{IL}^*_{set}$-model. There exists an $\mathbf{IL}$-model $M' = \langle W', R', S', \Vdash\rangle$ such that $W = W'$, $R = R'$ and for all formula's A and all $w \in W$ we have*

$$w \Vdash A \Leftrightarrow w \Vdash' A.$$

Proof. Put $W' = W$, $R' = R$ and

$$xS'_wy \Leftrightarrow xS_w\{y\}.$$

And let $\Vdash'$ agree with $\Vdash$ on propositional variables and extend it to more complex formulas as usual.

We first show that M' is an $\mathbf{IL}$-model.

1. R' is conversely well-founded since $R' = R$
2. R' is transitive since $R' = R$
3. If xS'_wy then $xS_w\{y\}$. So, $wR'x, y$ and thus wRx, y
4. If $wR'x$ then wRx. So, $yS_w\{y\}$ and thus yS'_wy
5. If $wR'xR'y$ then $wRxRy$. So, $xS_w\{y\}$ and thus xS'_wy
6. If xS'_wy and yS'_wz then $xS_w\{y\}$ and $yS_w\{z\}$. So, by 3.(c)', $xS_w\{z\}$ and thus xS'_wz

Secondly we show that if M is an $\mathbf{IL}^*_{\rm set}\mathsf{P}_0$-model, then M' is an $\mathbf{IL}\mathsf{P}_0$-model. Suppose $wR'xR'yS'_wy'R'z$. We have to show yS_xz. We have $wRxRyS_w\{y'\}$. Put $Z = \{z\}$. Then for all $s \in \{y\}$ there is some $t \in Z$ such that sRt. So by the P_0-condition (for $\mathbf{IL}^*_{\rm set}$-frames) we obtain some

$Z' \subseteq Z$ such that yS_xZ'. Clearly the only possibility for such Z' is Z. Thus, by definition of S', yS'_xz.

Finally we show that for all $w \in W$, M' and M force the same formulas A in w. Induction on A. If A is a propositional variable then this is clear by definition. Boolean connectives are trivial.

So, suppose $w \Vdash C \rhd D$. We show $w \Vdash' C \rhd D$. Pick x such that $wR'x$ and $x \Vdash' C$. By (IH), $x \Vdash C$. So, for some $Y \subseteq W$ we have xS_wY and $Y \Vdash D$. By Lemma 13.16, for some $y \in Y$, $xS_w\{y\}$. Fix such a y. By (IH) we have $y \Vdash' B$ and by definition of S', xS'_wy.

Next suppose, $w \Vdash' C \rhd B$. We show $w \Vdash C \rhd D$. Pick some x such that wRx and $x \Vdash C$. By (IH), $x \Vdash' C$. So, for some y we have $y \Vdash' D$ and xS'_wy. By definition of S', $xS_w\{y\}$. ⊣

13.2 The incompleteness of $\mathbf{IL}\mathsf{M_0P_0W}$

This section is primary devoted to showing a conjecture from [Joo98]. Namely that $\mathbf{IL}\mathsf{M_0P_0W}$ is an incomplete logic. To this end we use the model/bisimulation method from [Š91]. To start we introduce the schema R (which we have already touched upon in the previous section).

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

In what follows we will show that R is not derivable in $\mathbf{IL}\mathsf{M_0P_0W}$, but that it should have been were $\mathbf{IL}\mathsf{M_0P_0W}$ complete. As we will see, the characteristic Veltman frame class of $\mathbf{IL}\mathsf{R}$ is the same as the characteristic Veltman frame class of $\mathbf{IL}\mathsf{P_0}$. In the previous section we used generalized Veltman semantics to circumvent such situations, here we will work with Veltman models and bisimulation arguments.

Lemma 13.18. *Let $\mathbf{IL}'$ be some extension of $\mathbf{IL}$. Let N be a model such for all axioms X of $\mathbf{IL}'$ and any world $v \in N$ we have $v \Vdash X$. (We write $N \models \mathbf{IL}'$ for this.) Then for any theorem A of $\mathbf{IL}'$ we have $N \models A$.*

Proof. Induction on the derivation of A. If A is an axiom of **IL**' then $N \models A$ by assumption. First we show that validity on a model is preserved under generalization. Suppose $A = \Box A'$ and A' is valid on N. Pick any $x \in N$. To show that $x \Vdash A$ we need to show that for all y with xRy, $y \Vdash A'$. But by assumption A' holds everywhere on N so certainly on those y. Next we show that validity on a model is preserved under modes-ponens. Suppose $N \models A'$ and $N \models A' \to A$. Pick some x in N. Then $x \Vdash A'$ and $x \Vdash A' \to A$. Thus $x \Vdash A$. ⊣

Consider Figure 6. Let M be the smallest **IL** model based on this figure. So we take each S_v reflexive on $v\!\uparrow$, R and each S_v transitive and each S_v extends R on $v\!\uparrow$. w, $x_0, \ldots$, $y_0, \ldots$ are names for the worlds. a, b, c, d are proposition variables. They are placed exactly at the worlds in which they are true. (So, for example, in x_0 all of a, b, c and d are false.)

Lemma ($M \models \mathbf{IL}\mathsf{P_0}$). *For all formulas A and B and each $s \in M$,*

$$s \Vdash A \rhd \Diamond B \to \Box(A \rhd B).$$

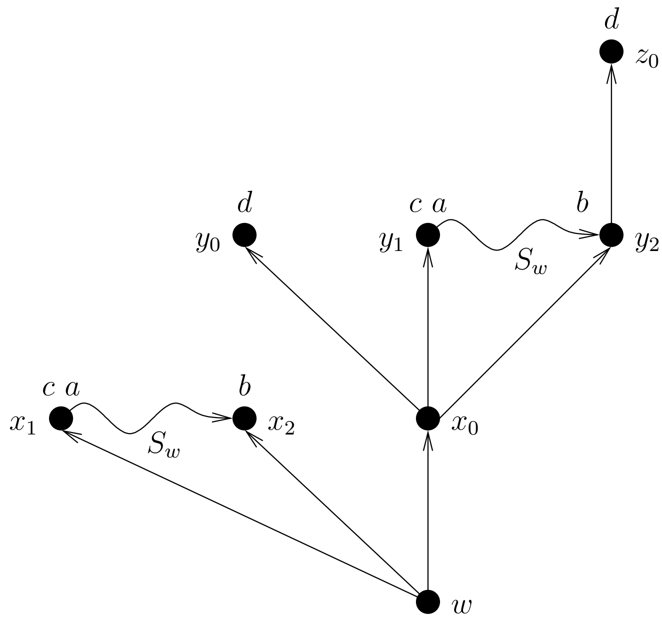

Figure 6: An **IL**-model M

Proof. Suppose $s \Vdash A \rhd \Diamond B$. Take some t for which sRt. It is sufficient to show that for all t', with tRt', $t' \nVdash A$. So take a t' with tRt'. Then $sRtRt'$ and thus $s = w$. Moreover $t = x_0$ or $t = y_2$.

case: $t = x_0$. Now $t' = y_0$ or $t' = y_1$.

case: $t' = y_0$. Clearly for no u, with $y_0 S_w u$, we have $u \Vdash \Diamond B$. So, since $w \Vdash A \rhd \Diamond B$, we have $y_0 \nVdash A$.

case: $t' = y_1$. If $y_1 \Vdash A$ then also $x_1 \Vdash A$. But for no u with $x_1 S_w u$, we have $u \Vdash \Diamond B$. In contradiction with $w \Vdash A \rhd \Diamond B$.

case: $t = y_2$. Now $t' = z_0$. Since $w \Vdash A \rhd \Diamond B$, $z_0 \nVdash A$.

$\dashv$

Lemma ($M \models \mathsf{M}_0$)**.** *For all formulas A, B and C and each $s \in M$.*

$$s \Vdash A \rhd B \to \Diamond A \wedge \Box C \rhd B \wedge \Box C.$$

Proof. Suppose $s \Vdash A \rhd B$. Let t be such that sRt and $t \Vdash \Diamond A \wedge \Box C$. We have to find an u with $tS_w u$ and $u \Vdash B \wedge \Box C$.

Since $t \Vdash \Diamond A$, $t = x_0$ or $t = y_2$. Moreover, in any case, $s = w$.

case: $t = x_0$. In this case $y_0 \Vdash A$ or $y_1 \Vdash A$.

case: $y_0 \Vdash A$. Since $w \Vdash A \rhd B$ we now must have that $y_0 \Vdash B$. So, since clearly $y_0 \Vdash \Box C$, we can take y_0 for u.

case: $y_1 \Vdash A$. Since $w \Vdash A \rhd B$ we now have $y_1 \Vdash B$ or $z_0 \Vdash B$ or $y_2 \Vdash B$. In all off these cases we can take y_1 resp. z_0 resp. y_2 for u. In the first two cases is this clear since trivially $y_1 \Vdash \Box C$ and $z_0 \Vdash \Box C$. That $y_2 \Vdash \Box C$ can be seen as follows. Since $x_0 \Vdash \Box C$ we have $y_0 \Vdash C$ and thus $z_0 \Vdash C$. But z_0 is the only R-successor of y_2. So, this implies, indeed, $y_2 \Vdash \Box C$.

case: $t = y_2$. In this case $z_0 \Vdash A$. Since $w \Vdash A \rhd B$, we have $z_0 \Vdash B$ and thus, since $z_0 \Vdash \Box C$, we can take z_0 for u.

⊣

Lemma ($M \models \mathsf{W}$). *For all formulas A and B and each $s \in W$,*

$$s \Vdash A \rhd B \to A \rhd B \wedge \Box\neg A.$$

Proof. It is well know that W is valid on any frame in which, for each w, $R \circ S_w$ is conversely well-founded. So it is certainly valid on models based on such frames. ⊣

Lemma ($M \not\models \mathsf{R}$).

$$w \not\Vdash a \rhd b \to \neg(a \rhd d) \wedge (\neg c \rhd d) \rhd b \wedge \Box c$$

Proof. $w \Vdash a \rhd b$ is clear. $x_0 \Vdash \neg c \rhd d$ is also clear. y_1 is an example for $x_0 \Vdash \neg(a \rhd d)$. But the only world in the model where $b \wedge \Box c$ holds is x_2. Which is not reachable from x_0 with an S_w relation. ⊣

Definition 13.19 (P_0-frame). We say that a frame is an P_0-frame iff. for all w, x, y, y', z for which $wRxRyS_wy'Rz$ we have yS_xz.

Lemma. *A frame $F = \langle W, R, S\rangle$ is a P_0-frame iff. $F \models \mathsf{P}_0$.*

Proof. ($\Rightarrow$) Suppose F is a P_0-frame. Let $\overline{F} = \langle W, R, S, \Vdash\rangle$ be a model based on F. Pick some $w \in W$. Suppose $w \Vdash A \rhd \Diamond B$. Pick some $x, y \in W$ such that $wRxRy$ and $y \Vdash A$. Then wRy and thus for some y' with yS_wy', we have $y' \Vdash \Diamond B$. So, there exists some z for which $y'Rz$ and $z \Vdash B$. Since F is a P_0-frame, yS_xz. We conclude $x \Vdash A \rhd B$.

($\Leftarrow$) Suppose $F \models \mathsf{P}_0$. Pick w, x, y, y', z in W for which $wRxRyS_wy'Rz$. Let p and q be distinct proposition variables. Define a model $\overline{F} = \langle W, R, S \Vdash\rangle$, based on F, as follows.

$$v \Vdash p \Leftrightarrow v = y$$
$$v \Vdash q \Leftrightarrow v = z$$

Since $F \models \mathsf{P}_0$, in particular $w \Vdash p \rhd \Diamond q \to \Box(p \rhd q)$. By definition of $\Vdash$, $w \Vdash p \rhd \Diamond q$ and thus $x \Vdash p \rhd q$. Since $y \Vdash p$ there must be some t for which yS_xt and $t \Vdash q$. By definition of $\Vdash$, the only candidate for this is z. ⊣

Lemma (Frame condition of R). *A frame F is a P_0-frame iff. $F \models \mathsf{R}$.*

Proof. ($\Rightarrow$) Suppose F is a P_0-frame. Let $\overline{F} = \langle W, R, S, \Vdash\rangle$ be a model based on F. Choose some $w \in W$ for which $w \Vdash A \rhd B$. Pick x such that wRx and $x \Vdash \neg(A \rhd D) \wedge (\neg C \rhd D)$. Let y be an example for $x \Vdash \neg(A \rhd D)$. e.g. xRy, $y \Vdash A$ but for no u with yS_xu we have $u \Vdash D$. There exists some y' with yS_wy' and $y' \Vdash B$. We will show that, additionally, $y' \Vdash \Box C$. (This will finish the proof since xRy implies xS_wy and thus we then have xS_wy'.) Suppose that this is not so. Take some z with $y'Rz$ and $z \Vdash \neg C$. By the P_0-condition, xRz and yS_xz. So, there exists some z' with zS_xz' and $z' \Vdash D$. But then also yS_xz'. In contradiction with the choice of y.

($\Leftarrow$) Assume $F \models \mathsf{R}$. This gives that for distinct proposition variables p, q, r, s the formula $p \rhd q \to \neg(p \rhd r) \wedge (\neg p \rhd r) \rhd q \wedge \Box s$ is valid in F. Now pick w, x, y, y', z in F and suppose $wRxRyS_wy'Rz$. Define a model $\overline{F} = \langle W, R, S, \Vdash$ as follows.

$$
\begin{aligned}
v \Vdash p &\Leftrightarrow v = y\\
v \Vdash q &\Leftrightarrow v = y'\\
v \Vdash r &\Leftrightarrow \neg yS_xv\\
v \Vdash s &\Leftrightarrow v \neq z
\end{aligned}
$$

Now $w \Vdash p \rhd q$ and thus $w \Vdash \neg(p \rhd r) \wedge (\neg p \rhd r) \rhd q \wedge \Box s$. Also $x \Vdash \neg(p \rhd r)$. Now suppose for a contradiction that not yS_xz. Then $z \Vdash r$ and thus $x \Vdash \neg p \rhd r$. But now we must be able to find some world z' for which $z' \Vdash q \wedge \Box s$. But the only world at which q holds is y'. And y' can reach z, at which $\neg s$ is true. A contradiction. $\dashv$

By the above lemma we can speek of R-frames instead of P_0-frames.

Corollary 13.20. $\mathbf{IL}\mathsf{M}_0\mathsf{P}_0\mathsf{W}$ *is incomplete.*

Proof. By the above lemma's and the fact that M is a W model. $\dashv$

Let us finish this section with a question.

Does there exists an extension of $\mathbf{IL}$, $\mathbf{IL}'$ say, such that $\mathbf{IL}' \not\vdash \mathsf{P}_0$ and $\mathbf{IL}'\mathsf{P}_0$ is complete?

14 Logics containing R

We start this section by showing that $\mathbf{IL}\mathsf{R}$ is stronger than $\mathbf{IL}\mathsf{M}_0\mathsf{P}_0$. In Corollary 13.20 we have already seen that $\mathbf{IL}\mathsf{M}_0\mathsf{P}_0$ does not prove R. To see that $\mathbf{IL}\mathsf{R}$ is stronger than $\mathbf{IL}\mathsf{M}_0\mathsf{P}_0$ we need to prove the following theorem.

Theorem 14.1. $\mathbf{IL}\mathsf{R} \vdash \mathsf{M}_0\mathsf{P}_0$

Proof. First we show $\mathbf{IL}\mathsf{R} \vdash \mathsf{M}_0$. Consider the following instantiation of R.

$$A \rhd B \to \neg(A \rhd \bot) \wedge (\neg C \rhd \bot) \rhd B \wedge \Box C \tag{35}$$

We have $\mathbf{IL} \vdash \neg(A \rhd \bot) \leftrightarrow \Diamond A$ and $\mathbf{IL} \vdash \neg C \rhd \bot \leftrightarrow \Box C$. So,

$$\mathbf{IL} \vdash (35) \leftrightarrow (A \rhd B \rightarrow \Diamond A \wedge \Box C \rhd B \wedge \Box C).$$

Now we will show $\mathbf{IL}\mathsf{R} \vdash \mathsf{P}_0$. Consider the following instantiation of R.

$$A \rhd \Diamond B \rightarrow \neg(A \rhd B) \wedge (B \rhd B) \rhd \Diamond B \wedge \Box\neg B \tag{36}$$

Clearly, $\mathbf{IL} \vdash (36) \leftrightarrow (A \rhd \Diamond B \rightarrow \neg(A \rhd B) \rhd \bot)$. For all formulas F, $\mathbf{IL} \vdash \Box F \leftrightarrow \neg F \rhd \bot$. So,

$$\mathbf{IL} \vdash (36) \leftrightarrow (A \rhd \Diamond B \rightarrow \Box(A \rhd B)).$$

$\dashv$

14.1 The Logic ILR

We will show that two principles that happen to have the same frame condition as R are provable in $\mathbf{IL}\mathsf{R}$.

Let us recall the scheme R.

$$A \rhd B \rightarrow \neg(A \rhd D) \wedge (\neg C \rhd D) \rhd B \wedge \Box C \tag{37}$$

Lemma 14.2.

1. $\mathbf{IL}\mathsf{R} \vdash p \rhd (q \wedge \Diamond r) \vee (q \wedge \Box s) \rightarrow \neg(p \rhd r) \rhd q \wedge \Box s$
2. $\mathbf{IL}\mathsf{R} \vdash p \rhd q \wedge \Box(\Box\neg r \rightarrow \Box s) \rightarrow \neg(p \rhd r) \rhd q \wedge \Box s$

Proof. We first show 1. Take $A = p$, $B = (q \wedge \Diamond r) \vee (q \wedge \Box s)$, $C = \neg r$, $D = r$ in (37) to obtain

$$\mathbf{IL}\mathsf{R} \vdash p \rhd (q \wedge \Diamond r) \vee (q \wedge \Box s) \rightarrow \neg(p \rhd r) \rhd ((q \wedge \Diamond r) \vee (q \wedge \Box s)) \wedge \Box\neg r. \tag{38}$$

We also have

$$\mathbf{IL}\mathsf{R} \vdash (((q \wedge \Diamond r) \vee (q \wedge \Box s)) \wedge \Box\neg r) \leftrightarrow (q \wedge \Box s). \tag{39}$$

Combining (38) with (39) we obtain 1. Note that $(q \wedge \Diamond r) \vee (q \wedge \Box s) \leftrightarrow q \wedge (\Box\neg r \rightarrow \Box s)$ to see some similarities with 2.

We now show 2. Take $A = p$, $B = q$, $C = \neg r$ and $D = r$ in (37) to obtain

$$\mathbf{IL}\mathsf{R} \vdash p \rhd q \rightarrow \neg(p \rhd r) \rhd q \wedge \Box\neg r. \tag{40}$$

We also have

$$\mathbf{IL}\mathsf{R} \vdash \Box(\Box\neg r \rightarrow \Box s) \rightarrow q \wedge \Box\neg r \rhd q \wedge \Box s. \tag{41}$$

Combining (40) with (41) we obtain 2. $\dashv$

14.2 The logic $\mathbf{IL}\mathsf{R}^*$

In a certain way one could see the W scheme as the interpretability form of Löb's axiom. We know that the principle R, in a sense, does not embody Löb's axiom. To put it more precisely, $\mathbf{IL}\mathsf{R} \nvdash \mathsf{W}$. There is an obvious way to incorporate W and R into one single principle.

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

Of course this principle is reminiscent of M_0^*.

Lemma 14.3. $\mathbf{IL}\mathsf{R}^* = \mathbf{IL}\mathsf{RW}$

Proof. It is obvious that $\mathbf{IL}\mathsf{R}^* \vdash \mathsf{R}$. By taking $F = \bot$ we see that $\mathbf{IL}\mathsf{R}^* \vdash \mathsf{M}_0^*$ and thus by Lemma 12.10 also $\mathbf{IL}\mathsf{R}^* \vdash \mathsf{W}$.

For the other direction we first apply W and then R. Thus, $A \rhd B \to A \rhd B \wedge \Box\neg A$, and $A \rhd B \wedge \Box\neg A \to \neg(A \rhd F) \wedge (\neg C \rhd F) \rhd B \wedge \Box\neg A \wedge \Box C$. $\dashv$

Before trying to give a modal completeness proof of $\mathbf{IL}\mathsf{R}^*$ it seems useful to submit it to some tests. If for some principle X we have $\mathbf{IL}\mathsf{R}^* \models \mathsf{X}$, we should try to find an $\mathbf{IL}\mathsf{R}^*$-proof of X. Two such examples are the following formulas.

- $A \rhd B \to B \wedge (\neg C \rhd F) \rhd B \wedge (A \vee \neg C \rhd F)$
- $A \rhd B \to B \wedge (\neg C \rhd A) \rhd B \wedge \Box C \wedge \Box\neg A$

14.3 Some remarks on modal completeness

The current challenge is of course, to decide whether $\mathbf{IL}\mathsf{R}$ and $\mathbf{IL}\mathsf{R}^*$ are modally complete. We conjecture that they both are indeed complete. We obtained a modal completeness result for $\mathbf{IL}\mathsf{W}^*$ by analyzing and adjusting the completeness proof for $\mathbf{IL}\mathsf{M}_0$. It is not likely that we can obtain a completeness proof for $\mathbf{IL}\mathsf{R}$ or $\mathbf{IL}\mathsf{R}^*$ by minor adjustments.

A completeness proof for $\mathbf{IL}\mathsf{R}$ or $\mathbf{IL}\mathsf{R}^*$ will however contain many of the ingredients that were already present in the completeness proof of $\mathbf{IL}\mathsf{M}_0$. Instead of the relation $\subseteq_\Box$ we shall need to work with a new relation $\subset_B$.

Definition 14.4. For Γ and Δ MCS's, we define $\Gamma \subset_B \Delta$ as $A \rhd B \in \Gamma \Rightarrow \Box\neg A \in \Delta$.

We have the following useful observations on this new relation.

- $\Gamma \subset_{B_0} \Delta \subset_{B_1} \Delta' \Rightarrow \Gamma \subset_{B_0} \Delta'$
- $\Gamma \subset_B \Delta \subseteq_\Box \Delta' \Rightarrow \Gamma \subset_B \Delta'$
- $\Gamma \subset_B \Delta \prec \Delta' \Rightarrow \Gamma \prec_B \Delta'$
- $\Gamma \subset_B \Delta \Rightarrow \Gamma \subseteq_\Box \Delta$

The principle R is tailored to prove the following lemma.

Lemma 14.5. *If $C \rhd D \in \Gamma \prec_{B_0} \Delta \prec_{B_1} \Delta' \ni C$, then there exists some Δ'' with $\Gamma \prec_{B_0} \Delta'' \ni D, \Box\neg D$ and $\Delta \subset_{B_1} \Delta''$.*

One of the difficulties with $\mathbf{IL}\mathsf{M}_0$ also arises in $\mathbf{IL}\mathsf{R}$. With respect to what B_1 should Lemma 14.5 be applied? In the $\mathbf{IL}\mathsf{M}_0$-case, this difficulty was taken care of by the invariant

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

In the case of $\mathbf{IL}\mathsf{R}$ we would like something similar. If $\{x_i\}_{1\leq i\leq n}$ is the set of worlds for which uRx_iRv, with $\nu(x_i) \prec_{B_i} \nu(v)$, then for some permutation π of the indices we should have

$$\nu(x_{\pi(1)}) \subset_{B_{\pi(1)}} \nu(x_{\pi(2)}) \ldots \subset_{B_{\pi(n-1)}} \nu(x_{\pi(n)}).$$

But, replacing only $\mathcal{I}_\Box$ can never be sufficient, as the frame condition of R is really different from that of M_0. A quasi-$\mathbf{IL}\mathsf{X}$-frame will always be defined so that its closure is an $\mathbf{IL}\mathsf{X}$-frame. In this closure process, things will drastically change due to more S_y-relations. This will make notions like the $\mathcal{N}_x^A$ (see Definition 11.8) completely useless.

One way to go about this, would be to introduce Σ_x-relations on quasi-$\mathbf{IL}\mathsf{R}$-frames that capture the dynamics of the S_x during the closure process. This is in complete analogy with the introduction of the K-relation (see Definition 11.5) to capture the dynamics of the R-relation in quasi-$\mathbf{IL}\mathsf{M}_0$-frames. Note that the K-relation can still be used in the model construction in the completeness proof of $\mathbf{IL}\mathsf{R}$.

15 Remarks on the interpretability logic of all reasonable arithmetical theories

Let us first recall Definition 2.28, that is, the definition of the interpretability logic of all reasonable arithmetical theories. We shall write $\mathbf{IL}$(All). We defined $\mathbf{IL}$(All) to be the set of modal formulas that are interpretability principles in any reasonable arithmetical theory. That is, the set of φ for which

$$\forall T \forall * \ T \vdash \varphi^*.$$

In [Vis88] $\mathbf{IL}$(All) was conjectured to be $\mathbf{IL}\mathsf{W}$. In [Vis91] this conjecture was falsified and strengthened to a new conjecture. It was now conjectured that $\mathbf{IL}\mathsf{W}^*$, which is a proper extension of $\mathbf{IL}\mathsf{W}$ is $\mathbf{IL}$(All).

In [Joo98] it was proved that the logic $\mathbf{IL}\mathsf{W}^*\mathsf{P}_0$ is a proper extension of $\mathbf{IL}\mathsf{W}^*$, and that $\mathbf{IL}\mathsf{W}^*\mathsf{P}_0$ is a subsystem of $\mathbf{IL}$(All). This falsified the conjecture from [Vis91]. In [Joo98] it is also conjectured that $\mathbf{IL}\mathsf{W}^*\mathsf{P}_0$ is not the same as $\mathbf{IL}$(All).

In [JV00] it is conjectured that $\mathbf{IL}\mathsf{W}^*\mathsf{P}_0 = \mathbf{IL}$(All). As we will see below we have that the logic $\mathbf{IL}\mathsf{R}^*$ is a subsystem of $\mathbf{IL}$(All) and a proper extension of $\mathbf{IL}\mathsf{W}^*\mathsf{P}_0$. This rejects the conjecture pronounced in [JV00]. With all this conjecturing and refuting of conjectures we are rather hesitant in proposing as a new conjecture that $\mathbf{IL}\mathsf{R} = \mathbf{IL}$(All).[27]

[27] Although we do have a certain hope.

15.1 Arithmetical Soundness of R

We shall give a proof that the new principle

$$\mathsf{R} := A \rhd B \to \neg(A \rhd D) \wedge (\neg C \rhd D) \rhd B \wedge \Box C$$

is arithmetically valid in all reasonable theories. In the proof we shall employ some well-known arithmetical facts. We will now first briefly summarize these facts.

Definition 15.1. A definable T-cut is a formula $I(x)$ with one free variable, such that $T \vdash I(0) \wedge \forall x\ (I(x) \to I(x+1))$. $\mathsf{Cut}(\cdot)$ will denote the function that assigns to the code of a formula φ, the code of the formula expressing that φ is a cut, that is, $\varphi(0) \wedge \forall x\ (\varphi(x) \to \varphi(x+1))$ (whenever φ is of the right format).

The function $\mathsf{Cut}(\cdot)$ is a very easy function. It is certainly provably total in $\mathsf{I}\Delta_0 + \Omega_1$. In this section we shall denote the translation of a formula φ under an interpretation j by $j(\varphi)$. If I is a cut and φ a formula, we shall by φ^I denote the formula φ, where all the quantifiers in φ are relativized to the cut I. The following lemma is mentioned (as an exercise) in [Pud85]. It is central to many arguments in the field of formalized interpretability.

Lemma 15.2 (Pudlák). *There exists a function f, provably total in $\mathsf{I}\Delta_0+\Omega_1$, such that for any reasonable arithmetical theory T, the following holds.*

$$T \vdash j : \alpha \rhd_T \beta \to [\Box_T \mathsf{Cut}(f(j)) \wedge \forall\, \sigma \in \Sigma_1!\ j : \alpha \wedge \sigma^{f(j)} \rhd_T \beta \wedge \sigma]$$

Another fact from arithmetic that we shall need, is that we can perform the Henkin construction using numbers from a cut. This is expressed by the following lemma.

Lemma 15.3. *For any reasonable arithmetical theory T we have that*

$$T \vdash \Box_T(\mathsf{Cut}(I)) \to \Diamond_T^I \alpha \rhd_T \alpha.$$

These two lemmas are enough to prove the arithmetical soundness of the principle R. Note that the j, I, α and β in Lemma 15.2 en 15.3 are parameters and hence could be universally quantified within the theory.

Theorem 15.4 (Soundness of R). *For any reasonable arithmetical theory T we have the following.*

$$T \vdash \alpha \rhd \beta \to \neg(\alpha \rhd \delta) \wedge (\neg\gamma \rhd \delta) \rhd \beta \wedge \Box\gamma$$

Proof. Let f denote the function from Lemma 15.2. To prove our theorem, we reason in T and assume $\alpha \rhd \beta$. Thus, for some interpretation j we have $j : \alpha \rhd \beta$. We now claim that

$$\neg(\alpha \rhd \delta) \wedge (\neg\gamma \rhd \delta) \to \Diamond(\alpha \wedge \Box^{f(j)}\gamma). \qquad (+)$$

Let us first see that this claim, indeed entails the result.

$$\begin{array}{rcl} \neg(\alpha \rhd \delta) \wedge (\neg\gamma \rhd \delta) & \rhd & \text{By } (+) \\ \Diamond(\alpha \wedge \Box^{f(j)}\gamma) & \rhd & \text{By } \mathsf{J5} \\ \alpha \wedge \Box^{f(j)}\gamma & \rhd & \text{By Lemma 15.2 and } j : \alpha \rhd \beta \\ \beta \wedge \Box\gamma & & \end{array}$$

Thus, now we only need to prove the claim. We will prove (+) by showing the logical equivalent

$$\Box(\alpha \to \Diamond^{f(j)}\neg\gamma) \wedge (\neg\gamma \rhd \delta) \to \alpha \rhd \delta \qquad (++)$$

We reason as follows.

$$\begin{array}{rcl} \Box(\alpha \to \Diamond^{f(j)}\neg\gamma) & \to & \text{By J1} \\ \alpha \rhd \Diamond^{f(j)}\neg\gamma & \to & \text{By Lemma 15.3 and J2} \\ \alpha \rhd \neg\gamma & & \end{array}$$

Combining this with $\neg\gamma \rhd \delta$ yields the required $\alpha \rhd \delta$. ⊣

15.2 Principles in ILP ∩ ILM

Both essentially reflexive theories and strong enough finitely axiomatized theories are certainly reasonable arithmetical theories. Consequently, **IL**(All) $\subseteq$ **IL**M ∩ **IL**P. Therefore, it is good to know which principles can be found in this intersection.

Definition 15.5. The logic **IL**[M ∨ P] is defined by the axiom scheme [M ∨ P] which is the following.

$$[\mathsf{M} \vee \mathsf{P}] := (A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C) \vee (E \rhd F \to \Box(E \rhd F))$$

Lemma 15.6. **IL**P ∩ **IL**M = **IL**[M ∨ P]

Proof. If **IL**[M ∨ P] $\vdash \varphi$, then **IL** $\vdash \alpha \to \varphi$ for some α that is a conjunction of instantiations of [M ∨ P]. Clearly **IL**M $\vdash \alpha$ & **IL**P $\vdash \alpha$, whence $\varphi \in$ **IL**P ∩ **IL**M.

If, on the other hand, for some φ both **IL**M $\vdash \varphi$ and **IL**P $\vdash \varphi$, we can reason as follows. For some conjunction α of instantiations of the M-axiom scheme, we get

$$\mathbf{IL} \vdash \alpha \to \varphi.$$

Likewise, we get for some β, which is the conjunction of P-instantiations, that **IL** $\vdash \beta \to \varphi$. Consequently

$$\mathbf{IL} \vdash \alpha \vee \beta \to \varphi.$$

We can rewrite $\alpha \vee \beta$ as a conjunction of disjunctions by distributing the $\vee$ at the top level in $\alpha \vee \beta$ over all the conjunctions at the top level in α and in β. This yields an equivalent formula γ which is a conjunction of [M ∨ P]-instantiations. We conclude that **IL**[M ∨ P] $\vdash \varphi$. ⊣

It is clear that Lemma 15.6 is neither a manageable nor an informative characterization of **IL**M ∩ **IL**P. The question arises whether such a characterization does exist. Such a characterization should speed up the process of judging candidates for new principles. For example, $A \rhd B \to \Diamond(A \wedge \Box C) \rhd B \wedge \Box C$ proves all of M_0, P_0 and W, but is not provable in **IL**P.

The following items are all somehow related to **IL**(All).

- The principle $A \rhd B \to \Diamond(A \wedge \Box(A \vee B \to \Box C)) \rhd B \wedge \Box C$ and other principles related to $xRyRzS_xu \to yRu \vee \forall v\ (uRv \to zRv)$.
- Principles related to $xRyRzS_xuRv \to (yRu \vee zRv)$, like $A \rhd B \wedge \Diamond C \to (\Diamond A \to (\Diamond(A \wedge \Diamond C) \vee \Diamond(B \wedge \Diamond C)))$.
- Principles related to $xRyRzS_xuRv \to (zS_yu \vee zRv)$.
- Principles like $(\Diamond A \rhd \Diamond B) \wedge (A \rhd B) \to \Box(\Diamond A \to \Diamond B)$.
- Principles like $A \rhd B \to \Diamond A \wedge \Box\Box C \rhd B \wedge \Box C$, that are related to $xRyRzS_xuRv \to \exists w\ yRwRv$.

At this stage, both positive and negative information will be very informative.

16 Concluding

What are we to conclude at the end of this paper? We have presented a method for completeness proofs. And indeed the results are promising. In future completeness proofs, large parts of work can be skipped by referring to the results in this paper. So, in this light our mission has not been in vain. We do think however that the presentation is not yet optimal. Nevertheless we have gained some insights in where the bottle-neck of the completeness proofs can be expected.

16.1 Future research

Now that we have developed such precise and powerful machinery, new arithmetically interesting questions can be addressed. We mention two of them.

- The interpretability logic of primitive recursive arithmetic. In [Joo03] the modal study of this logic has been initiated. Without a concise modal tool box such a study seems inconceivable.
- The interpretability logic of all reasonable arithmetical theories still remains uncovered. In this project the first step to take would be to prove our conjecture on the modal completeness of $\mathbf{IL}\mathsf{R}^*$.

The harvest has only just begun!

16.2 Refining techniques

One of our goals, was to isolate a part of work that is present in all modal completeness proofs, and deal with it once and for all. In realizing this goal we were driven by two opposed forces. On the one hand, generalicity whence a wide scope of applicability, on the other hand, specificallity as to save as much work for the applications you are thinking of.

When we started this project we tried to aim at generalicity. It turned out that interpretability logics can have wild semantical peculiarities which are hard to catch in a general approach.

For example, in the definition of quasi-$\mathbf{IL}\mathsf{X}$-frames, we tried to be as general as possible. That is to say, we tried to specify the widest class of structures that can be closed to an **IL**X-frame. This is an interesting question, but not the question we started out with. We are interested in completeness.

By taking such a general notion of **IL**X-frame, our invariants should be as strong as to single out in this large class, the structures which can be endowed with a truth lemma. But, if we look at the "trajectory" of a single pointed model (in completeness proofs, you always start out with such a model) throughout the construction, we see that we can actually employ a notion of quasi-**IL**X-frame that is a lot more specific. Therefore, we think that a lot can be gained by using stricter notions of quasi-ness.[28] In general, finding a better balance between generality and easy applicability to our purposes seems desirable.

Furthermore, we think that extending the notion of criticality as mentioned in Section 11.8 looks very promising in the way it dealt with the difficulty of the transitivity of the S_x-relations in **IL**M_0.

We think that somehow completeness proofs in interpretability logics are intrinsically elaborative. This is not very pleasant, especially if one works in the area of interpretability logics. One could wonder whether we have the right notion of modal semantics. We would be more than happy if an easier notion would show up. However we should keep in mind that the $\rhd$-modality is used to model a Σ_3-complete notion. Viewed in this light it seems miraculous that there even exists a decidable semantics.

The current semantics has also proven to be a good one with respect to the mathematical entity that it is modeling. Two principles, P_0 and R, which are valid in all reasonable arithmetical theories, have been discovered on the basis of considerations on this semantics only.

The construction method we have presented here could also be applied to get some results on complexity matters in interpretability logics. In such an analysis one should study strategies for efficiently eliminating problems and deficiencies. One could think of the following ingredients. First eliminating problems. Like this some deficiencies might be dealt with automatically. But also one should think on the order in which problems are eliminated.

Acknowledgment Thanks to: Lev, Dick, Albert, Roos, Yde, Panikowski, Remco, Vincent, Merlino

[28]We have even considered to build some path coding in to the notion of quasi-**IL**X-frames.

References

[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.

[Bus98] S.R. Buss. Introduction to Proof Theory. In S.R. Buss, editor, *Handbook of Proof Theory*, pages 1–78, Amsterdam, 1998. Elsevier, North-Holland.

[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.

[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.

[Joo03] J.J. Joosten. The closed fragment of the interpretability logic of PRA with a constant for $\mathrm{I}\Sigma_1$. Logic Group Preprint Series 128, University of Utrecht, February 2003.

[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.

[Pud85] P. Pudlák. Cuts, consistency statements and interpretations. *Journal of Symbolic Logic*, 50:423–441, 1985.

[Sha88] V. Shavrukov. The logic of relative interpretability over Peano arithmetic (in Russian). Technical Report Report No.5, Steklov Mathematical Institute, Moscow, 1988.

[Sha97] V.Yu. Shavrukov. Interpreting reflexive theories in finitely many axioms. *Fundamenta Mathematicae*, 152:99–116, 1997.

[Sol76] R.M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 28:33–71, 1976.

[Š91] V. Švejdar. Some independence results in interpretability logic. *Studia Logica*, 50:29–38, 1991.

[Vis88] A. Visser. Preliminary notes on interpretability logic. Technical Report LGPS 29, Department of Philosophy, Utrecht University, 1988.

[Vis90] A. Visser. Interpretability logic. In P.P. Petkov, editor, *Mathematical Logic*, pages 175–208. Plenum Press, New York, 1990.

[Vis91] A. Visser. The formalization of interpretability. *Studia Logica*, 50(1):81–106, 1991.

[Vis95] A. Visser. A course on bimodal provability logic. *Annals of Pure and Applied Logic*, pages 109–142, 1995.

[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.