Modal Matters for Interpretability Logics

Evan Goris, *Department of Computer Science, the Graduate Center of the City University of New York, 365 Fifth avenue, New York, NY 10016, USA, E-mail: evangoris@gmail.com*

Joost J. Joosten, *Mathematical Institute of the Academy of Sciences of the Czech Republic, Zitna 25, CZ-11567, Praha 1, Prague, Czech Republic, E-mail: jjoosten@phil.uu.nl*

Abstract

This paper is the first in a series of three related papers on modal methods in interpretability logics and applications. In this first paper the fundaments are laid for later results. These fundaments consist of a thorough treatment of a construction method to obtain modal models. This construction method is used to reprove some known results in the area of interpretability like the modal completeness of the logic **IL**.

Next, the method is applied to obtain new results: the modal completeness of the logic $\mathbf{IL}\mathsf{M}_0$, and modal completeness of $\mathbf{IL}\mathsf{W}^*$.

Keywords: Interpretability Logics, Modal Logics, Modal Completeness

1 Introduction

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

The notion of interpretability that we are primarily interested in, is the notion of relativized interpretability as studied e.g. by Tarski et al in [25]. Roughly, a theory U interprets a theory V –we write $U \rhd V$– if U proves all theorems of V under some structure preserving translation. We allow for relativization of quantifiers. It is defendable to say that U is as least as strong as V if $U \rhd V$. We think that it is clear that interpretations are worth to be studied, as they are omnipresent in both mathematics and meta-mathematics (Langlands Program, relative consistency proofs, undecidability results, Hilberts Programme and so forth).

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

The main aim of this paper is to establish some modal techniques/toolkit for interpretability logics. Most techniques are aimed at establishing modal completeness results. As we shall see, in the field of interpretability logics, modal completeness can be a sticky business compared to unary modal logics. In this paper we make a first attempt at pulling some (more) thorns out. Significant progress with this respect has also been made by de Jongh and Veltman [8]. In [18], a masters thesis written under the supervision of de Jongh, the current version of the construction method was already present in a rudimentary form.

We have a feeling that the general modal theory of interpretability logics is getting more and more mature. For example, fixed point phenomena and interpolation are quite well understood ([10], [1], [31]).

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

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

2 Interpretability logics

In this section we will define the basic notions that are needed throughout the paper. We advise the reader to just skim through this section and use it to look up definitions whenever they are used in the rest of the paper.

2.1 Syntax and conventions

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

J5 $\Diamond A \rhd A$

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

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

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

Lemma 2.2
1. $\mathbf{IL} \vdash \Box A \leftrightarrow \neg A \rhd \bot$
2. $\mathbf{IL} \vdash A \rhd A \wedge \Box\neg A$
3. $\mathbf{IL} \vdash A \vee \Diamond A \rhd A$

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

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

M $A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C$
P $A \rhd B \to \Box(A \rhd B)$
$\mathsf{M_0}$ $A \rhd B \to \Diamond A \wedge \Box C \rhd B \wedge \Box C$
W $A \rhd B \to A \rhd B \wedge \Box\neg A$
$\mathsf{W^*}$ $A \rhd B \to B \wedge \Box C \rhd B \wedge \Box C \wedge \Box\neg A$
$\mathsf{P_0}$ $A \rhd \Diamond B \to \Box(A \rhd B)$
R $A \rhd B \to \neg(A \rhd \neg C) \rhd B \wedge \Box C$

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

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

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

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

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

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

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

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, the *successor* relation $\prec$, the *C-critical successor* relation $\prec_C$ and the *Box-inclusion* relation $\subseteq_\Box$.

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

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

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

LEMMA 2.8
Let Γ and Δ denote maximal $\mathbf{IL}\mathsf{X}$-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 $\mathbf{IL}$-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. $uS_xvS_xw \rightarrow uS_xw$

$\mathbf{IL}$-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.

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

Definition 2.10

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 $p \in \mathsf{Prop}$
- $w \widetilde{\nVdash} \bot$
- $w \widetilde{\Vdash} A \to B$ 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 $\forall u\ (wRu \wedge u \widetilde{\Vdash} A \Rightarrow \exists v\ (uS_w v \widetilde{\Vdash} B))$

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

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

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

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

Definition 2.11 (Generated Submodel)

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

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

In other words, $M{\upharpoonright}m$ is simply the restriction of M to $m{\upharpoonright}{*}$.

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

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

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

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

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

If $F = \langle W, R, S\rangle$ is an **IL**-frame and $\Vdash$ a subset of $W \times \mathsf{Prop}$, we denote by $\langle W, \Vdash\rangle$ the **IL**-model that is based on F and has forcing relation $\Vdash$. We say that a frame F

makes a formula φ true, and write $F \models \varphi$, if any model based on F makes φ true. In a second-order formula:

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

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

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

for frames F, and

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

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

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

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

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

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

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

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

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

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

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

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

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

---

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

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

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

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

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

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

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

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

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

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

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

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

PROOF. By induction on the derivation of φ. ∎

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

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

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

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

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

PROOF. Trivial. By the deduction theorem $\Gamma \vdash_{\mathbf{IL}\mathsf{X}} \varphi \Leftrightarrow \vdash_{\mathbf{IL}\mathsf{X}} \bigwedge \Gamma \rightarrow \varphi$. 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 [13]. This will allow us to do reasonable coding of syntax. We call these theories reasonable arithmetical theories.

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

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

For a technically more transparent definition we will treat functions as relation symbols for which the functionality axiom should hold.

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 $(\bot)^j = (\vee_\varnothing)^j = \vee_\varnothing = \bot$
- $(\forall x\ \varphi)^j = \forall x\ (\delta(x) \to \varphi^j)$
- $V \vdash \varphi \Rightarrow U \vdash \varphi^j$

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

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

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

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

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

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

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

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

Theorem 2.24 (Solovay [24])
$\mathbf{PL}(\mathrm{T}) = \mathbf{GL}$ 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 [3], Shavrukov [23])
If T is an essentially reflexive theory, then $\mathbf{IL}(\mathrm{T}) = \mathbf{IL}\mathsf{M}$.

Theorem 2.27 (Visser [28])
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 [19] 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 the sequel to this paper.

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 $\Sigma_n!$-formulas instead of Σ_n-formulas and Σ_n-formulas instead of $\Sigma_n(T)$-formulas.

3 General exposition of the construction method

A central result in this paper is given by a construction method that shall be worked out in the next section. Most of the applications of this 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 of $\mathbf{IL}\mathsf{X}$ 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 shall lay out the basic ideas which are involved in the construction method. In particular, 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 $\varphi \in \nu(x)$.

[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.

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 **GL**-consistent set, as any finite part of it is satisfiable in some world in some model. However, it is not hard to see that in no **IL**-model all of Γ can hold simultaneously in some world in it.

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

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

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

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. An examples of this sort of calculation is Lemma Lemma 6.19.

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.

[6] This example comes from Fine and Rautenberg and is treated in Chapter 7 of [5].

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 $|\!\!\!\sim$ 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.

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

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

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

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 $x|\!\!\!\not\sim \neg(\varphi \rhd \psi)$. We define a deficiency to be a configuration such that $\varphi \rhd \psi \in x \cap \mathcal{D}$ but $x|\!\!\!\not\sim \varphi \rhd \psi$. 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 $yS_x z$, we have $\psi \in z$. This can simply be eliminated by adding a z with $yS_x z$ 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, $yS_x z \Rightarrow z = y$ and thus indeed, it is not possible to reach a world where ψ holds. Now here is one complication.

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

With the elimination of a problem, things are more subtle. When we introduced $y \ni \varphi, \neg\psi$ to eliminate a problem $\neg(\varphi \rhd \psi) \in x \cap \mathcal{D}$, we did indeed eliminate it, as for no z with $yS_x z$ 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 [6]). 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 [4]). Or one can mention uniform methods like the use of Shalqvist formulas or the David Lewis theorem [5]. A very secure method is to construct counter models piece by piece. A nice example can be found in [5], Chapter 10. In [15] and in [14] 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 [6]). However, for logics that contain **IL**$\mathsf{M_0}$ but not **IL**M it is completely unclear how to obtain completeness via an unraveling and we are forced into more secure methods like the above mentioned building of models piece by piece. And this is precisely what we do in this paper.

Decidability and the finite model property are two related issues that more or less seem to divide the landscape of interpretability logics into the same classes. That is, the proof that **IL** has the finite model property is relatively easy. The same can be said about **IL**M. For logics like **IL**$\mathsf{M_0}$ the issue seems much more involved and a proper proof of the finite model property, if one exists at all, has not been given yet. Alternatively, one could resort to other methods for showing decidability like the Mosaic method [4].

4 The construction method

In this section we describe our construction method in full detail. Later sections are applications of the construction method.

4.1 Preparing the construction

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

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

If there is no chance of confusion we will just speak of labeled frames or even just of frames rather than **IL**X-labeled frames. Labeled frames inherit all the terminology and notation from normal frames. Note that an **IL**X-labeled frame need not be, and shall in general not be, an **IL**X-frame. If we speak about a labeled **IL**X-frame we always mean an **IL**X-labeled **IL**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 $(W \cup (W \times W)) \times (\mathsf{Form}_{\mathbf{IL}} \cup \{\Gamma \mid \Gamma \text{ is a maximal } \mathbf{IL}\mathsf{X} \text{ consistent set}\})$ such that the following properties hold.

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

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 $\mathsf{Form}_{\mathbf{IL}} \cup \{\Gamma \mid \Gamma \text{ is a maximal } \mathbf{IL}\mathsf{X} \text{ consistent set}\}$. Thus, from now on, we shall stick to the notation $\nu(x, y) = C$.

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

We sometimes refer to $\overline{F}$ as the model induced by the frame F. Alternatively we will speak about the model corresponding to F. Note that for **IL**X-models M, we have $\overline{\overline{M}} = M$, but in general[8] $\overline{\overline{F}} \neq F$ for **IL**X-labeled frames F.

The following definition is tailored to follow C-critical successors that we will introduce during our construction process. The critical cones will be used to guarantee that some world that was introduced to be a C-critical successor, will always remain C-critical.

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

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

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

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

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

We also note that there is some redundancy in Definitions 4.3 and 4.4. The last clause in the inductive definitions demands closure of the cone under R-successors.

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

[8] It is easily seen that the reason for this inequality is the fourth requirement in our formal definition of ν.

But from Definition 2.9.5 closure of the cone under R follows from closure of the cone under S_x. We have chosen to explicitly adopt the closure under the R. In doing so, we obtain a notion that serves us also in the environment of so-called quasi frames (see Definition 5.1) in which not necessarily $(x\upharpoonright)^2 \cap R \subseteq S_x$.

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

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

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

DEFINITION 4.6 (Temporary definition)
[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 4.7 (Deficiencies)
Let $\mathcal{D}$ be some set of sentences and let $F = \langle W, R, S, \nu \rangle$ be an **IL**X-labeled frame. A *$\mathcal{D}$-deficiency* is a triple $\langle x, y, C \rhd D\rangle$ with xRy, $C \rhd D \in \nu(x) \cap \mathcal{D}$, and $C \in \nu(y)$, but for no z with yS_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 4.8
Let A be a formula. We define the *single negation* of A, we write ${\sim}A$, as follows. If A is of the form $\neg B$ we define ${\sim}A$ to be B. If A is not a negated formula we set ${\sim}A := \neg A$.

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

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

PROOF. The proof is really very simple and precisely shows the interplay between all the ingredients. ∎

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

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

1. $xRy \Rightarrow \nu(x) \prec \nu(y)$

[9] We will eventually work with Definition 4.11.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4.2 The main lemma

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

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

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

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

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

In case such a set of invariants $\mathcal{I}$ exists, we have that any **IL**X-labeled adequate frame F satisfying all the invariants can be extended to some labeled adequate **IL**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 **IL**X is a complete logic.

PROOF. We shall first give a short version of the proof which is more readable. However, as this lemma is key to this paper, we shall also present a version of the proof including more detail.

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

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

But this just follows from the above. If **IL**X $\nvdash A$, we can find a maximal **IL**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 **IL**X-frame $\hat{F}$ on which a truth lemma with respect to $\mathcal{D}$ holds. Thus certainly $\hat{F}, x \Vdash \neg A$, that is, A is not valid on the model induced by $\hat{F}$. This ends the short version of the proof.

More detailed version of the proof: So, let **IL**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 **IL**X-frame $\hat{F}$ on which a truth lemma w.r.t. $\mathcal{D}$ holds.

In the light of Lemma 4.9 and of Remark 4.12 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 $X := \{x_0, x_1, \ldots\}$ 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[10] $x \prec_B z$ 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 $xRy \rightarrow \exists\, A{\in}(\nu(y) \setminus \nu(x)) \cap \{\Box\neg D \mid D$ a subformula of some $B \in \mathcal{D}\}$. Clearly, if in some F_i we have that $x_0Rx_1R\ldots Rx_m$ 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 such $\mathcal{D}$, all the corresponding invariants $\mathcal{I}$ hold on any one-point labeled frame, we are to see that **IL**X is a complete logic, that is, **IL**$\mathsf{X} \nvdash A \Rightarrow \exists M\ (M \models X\ \&\ M \models \neg A)$. But this just follows from the above. If **IL**$\mathsf{X} \nvdash A$, we can find a maximal **IL**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 **IL**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.

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

Lemma 4.17

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

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

Proof. So, consider $\neg(A \rhd B) \in \Gamma$, and suppose that no required Δ exists. We can then find a[11] formula C for which $C \rhd B \in \Gamma$ such that

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

Consequently

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

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

For deficiencies there is a similar lemma.

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

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

4.3 Completeness and the main lemma

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

Remark 4.19
A modal completeness proof for an interpretability logic $\mathbf{IL}\mathsf{X}$ 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.

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

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

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

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

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

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

5 The logic IL

The modal logic **IL** has been proved to be modally complete in [8]. 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. $yS_xz \to xRy$ & 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. $y{\in}\mathcal{C}_x^A \to \nu(x) \prec_A \nu(y)$

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

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

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

(i) $\gamma = \langle 0, a, b, c\rangle$ with $F_n \models aRbRc$ but $F_n \not\models aRc$
(ii) $\gamma = \langle 1, a, b\rangle$ with $F_n \models aRb$ but $F_n \not\models bS_ab$
(iii) $\gamma = \langle 2, a, b, c, d\rangle$ with $F_n \models bS_acS_ad$ but not $F_n \models bS_ad$
(iv) $\gamma = \langle 3, a, b, c\rangle$ with $F_n \models aRbRc$ but $F_n \not\models bS_ac$

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

Let $<_0$ be the well-ordering on

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

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

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

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

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

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

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

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

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

∎

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

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

holds. Then, 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 $aRbRc \rightarrow \nu(a) \subseteq_\Box \nu(c)$. Thus, if $A\in((\nu(c) \setminus \nu(b)) \cap \{\Box D \mid D \in \mathcal{D}\})$, then certainly $A \notin \nu(a)$. ∎

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

5.2 Modal completeness

Theorem 5.4
IL is a complete logic.

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

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

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

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

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

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

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

   It is easy to see that G' is actually a quasi-frame. Note that if $G' \models xRb$, then x must be a and whence $\nu(x) \prec \nu(b)$ by definition of $\nu(b)$. Also it is not hard to see that if $b \in \mathcal{C}_x^C$ for $x{\neq}a$, that then $\nu(x) \prec_C \nu(b)$. For, $b \in \mathcal{C}_x^C$ implies $a \in \mathcal{C}_x^C$ whence $\nu(x) \prec_C \nu(a)$. By $\nu(a) \prec \nu(b)$ we get that $\nu(x) \prec_C \nu(b)$. In case $x{=}a$ we see that by definition $b \in \mathcal{C}_a^B$. But, we have chosen Δ so that $\nu(a) \prec_B \nu(b)$. We also see that G' satisfies the invariant as $\Box\neg A \in \nu(b) \setminus \nu(a)$ and $\sim A \in \mathcal{D}$.
2. With Lemma 5.2 we extend G' to an adequate labeled **IL**-frame G. Corollary 5.3 tells us that the invariant indeed holds at G. Clearly $\langle a, \neg(A \rhd B)\rangle$ is no longer a problem in G.

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

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

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

   Again it is not hard to see that G' is a quasi-frame that satisfies the invariant. Clearly R is conversely well-founded. The only new S in G' is bS_ac, but we also defined aRc. We have chosen Δ' such that $\nu(a) \prec_B \nu(c)$. Clearly $\Box\neg D \notin \nu(a)$.
2. Again, G' is closed off under the frame conditions with Lemma 5.2. Again we note that the invariant is preserved in this process. Clearly $\langle a, b, C \rhd D\rangle$ is not a deficiency in G.

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

∎

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 differ from logic to logic. 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 $\mathbf{IL}\mathsf{M}_0$

This section is devoted to showing the following theorem.

Theorem 6.1
$\mathbf{IL}\mathsf{M}_0$ is a complete logic.

It turns out that the modal frame condition of $\mathbf{IL}\mathsf{M}_0$ gives rise to a bewildering structure of possible models that seems very hard to tame. As M_0 is in $\mathbf{IL}$(All), it is important that the class of $\mathbf{IL}\mathsf{M}_0$-frames is well understood. For a long time $\mathbf{IL}\mathsf{W}^*$ has been conjectured ([29]) to be $\mathbf{IL}$(All). A first step in proving this conjecture would have been a modal completeness proof of $\mathbf{IL}\mathsf{W}^*$.

It is well known that $\mathbf{IL}\mathsf{W}^*$ is the union of $\mathbf{IL}\mathsf{W}$ and $\mathbf{IL}\mathsf{M}_0$, see Lemma 7.3. The modal completeness of $\mathbf{IL}\mathsf{W}$ was proved in [9]. So, the missing link was a modal completeness proof for $\mathbf{IL}\mathsf{M}_0$. In [18] a proof sketch of this completeness result was given. In this paper we give for the first time a fully detailed proof.

In the light of Remark 4.19 a proof of Theorem 6.1 boils down to giving the four ingredients mentioned there. Sections 6.3, 6.4, 6.5, 6.6 and 6.7 below contain those ingredients. Before these main sections, we have in Section 6.2 some preliminaries. We start in Section 6.1 with an overview of the difficulties we encounter during the application of the construction method to $\mathbf{IL}\mathsf{M}_0$.

6.1 Overview of difficulties

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

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

Item 1. does not impose any direct difficulties. But some do emerge when we try to deal with the difficulties concerning Item 2. So let us see why it is difficult to ensure 2. Suppose we have $wRxRyS_wy'Rz$. The M_0–frame condition (Theorem 6.22) requires that we also have xRz. So, from 2. and the M_0–frame condition we obtain $wRxRyS_wy'Rz \rightarrow \nu(x) \prec \nu(z)$. A sufficient (and in certain sense necessary) condition is,

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

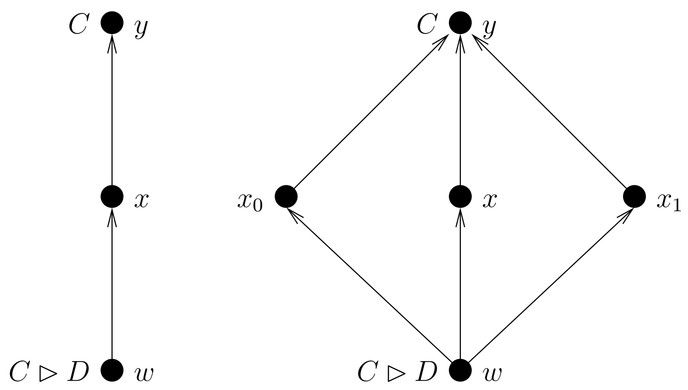

FIG. 1. A deficiency in w w.r.t. y

Let us illustrate some difficulties concerning this condition by some examples. Consider the left model in Figure 1. That is, we have a deficiency in w w.r.t. y. Namely, $C \rhd D \in \nu(w)$ and $C \in \nu(y)$. If we solve this deficiency by adding a world y', we thus require that for all x such that $wRxRy$ we have $\nu(x) \subseteq_{\Box} \nu(y')$. This difficulty is partially handled by Lemma 6.2 below. We omit a proof, but it can easily be given by replacing in the corresponding lemma for **IL**M, applications of the M-axiom by applications of the M_0-axiom.

LEMMA 6.2
Let Γ, Δ be MCS's such that $C \rhd D \in \Gamma$, $\Gamma \prec_A \Delta$ and $\Diamond C \in \Delta$. Then there exists some Δ' with $\Gamma \prec_A \Delta'$, $\Box\neg D, D \in \Delta'$ and $\Delta \subseteq_{\Box} \Delta'$.

Let us now consider the right most model in Figure 1. We have at least for two different worlds x, say x_0 and x_1, that $wRxRy$. Lemma 6.2 is applicable to $\nu(x_0)$ and $\nu(x_1)$ separately but not simultaneously. In other words we find y'_0 and y'_1 such that $\nu(x_0) \subseteq_{\Box} \nu(y'_0)$ and $\nu(x_1) \subseteq_{\Box} \nu(y'_1)$. But we actually want one single y' such that $\nu(x_0) \subseteq_{\Box} \nu(y')$ and $\nu(x_1) \subseteq_{\Box} \nu(y')$. We shall handle this difficulty by ensuring that it is enough to consider only one of the worlds in between w and y. To be precise, we shall ensure $\nu(x') \subseteq_{\Box} \nu(x)$ or $\nu(x) \subseteq_{\Box} \nu(x')$.

But now some difficulties concerning Item 1. occur. In the situations in Figure 1 we were asked to solve a deficiency in w w.r.t. y. As usual, if $w \prec_A y$ then we should be ably to choose a solution y' such that $w \prec_A y'$. But Lemma 6.2 takes only criticallity of x w.r.t. w into account. This issue is solved 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 2. That is, we have a deficiency in w w.r.t. y'. Namely, $C \rhd D \in \nu(w)$ and $C \in \nu(y')$. If we add a world y'' to solve this deficiency, as in the middle model, then by transitivity of S_w we have yS_wy'', as shown in the rightmost model. So, we require that $\nu(x) \subseteq_{\Box} \nu(y'')$. But we might very well have $\Diamond C \notin \nu(x)$. So the Lemma 6.2 is not applicable.

In Lemma 6.19 we formulate and prove a more complicated version of the Lemma 6.2 which basically says that if we have chosen $\nu(y')$ appropriately, then we can choose $\nu(y'')$ such that $\nu(x) \subseteq_{\Box} \nu(y'')$. And moreover, Lemma 6.19 ensures us that we can,

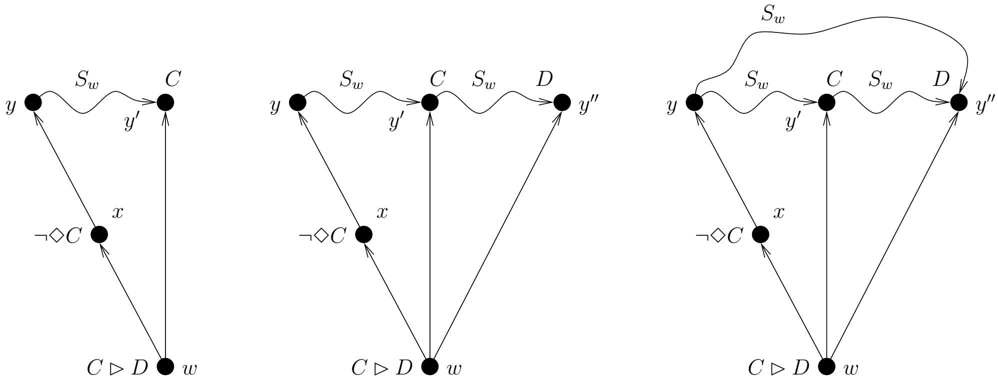

FIG. 2. A deficiency in w w.r.t. y'

indeed, choose $\nu(y')$ appropriate.

6.2 Preliminaries

DEFINITION 6.3 (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 fix the following fairly standard notations. T^{tr} is the transitive closure of T; T^* is the transitive reflexive closure of T; $xT;T'y \Leftrightarrow \exists t\ xTtT'y$; $xT^1y \Leftrightarrow xTy \wedge \neg\exists t\ xTtTy$; $xT^{\geq 2}y \Leftrightarrow xTy \wedge \neg(xT^1y)$ and $xT \cup T'y \Leftrightarrow xTy \vee xT'y$.

DEFINITION 6.4 ($\mathcal{S}_w$)
Let $F = \langle W, R, S, \nu\rangle$ be a quasi–frame. For each $w \in W$ we define the relation $\mathcal{S}_w$, of pure S_w transitions, as follows.

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

DEFINITION 6.5 (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.[12]

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 6.10 below.) We would like to predict on forehand which extra R relations will be added during this construction. The following definition does just that.

[12] 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.

Definition 6.6 ($K(F)$, K)
Let $F = \langle W, R, S, \nu \rangle$ be a quasi–frame. We define $K = K(F)$ to be the smallest binary relation on W such that

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

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

Lemma 6.7
Let $F_0 = \langle W, R_0, S, \nu \rangle$ and $F_1 = \langle W, R_1, S, \nu \rangle$ be quasi–frames. If $R_1 \subseteq K(F_0)$ and $R_0 \subseteq K(F_1)$. Then $K(F_0) = K(F_1)$.

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

Lemma 6.8 (Calculation of K)
Let $F = \langle W, R, S, \nu \rangle$ be a quasi–frame. Let $K = K(F)$ and suppose K conversely well–founded. Let T be a binary relation on W such that

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

Then we have the following.

(a) $K = T^{\mathrm{tr}}$
(b) $xK^1y \to xTy$

Proof. To see (a), it is enough to see that T^{tr} satisfies the three properties of the definition of K (Definition 6.6). Item (b) follows from (a). ∎

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

Definition 6.9 ($\mathcal{N}_w^C$)
For any quasi–frame F we define $\mathcal{N}_w^C$ to be the smallest set such that

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

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

Definition 6.10 (Quasi–$\mathbf{IL}\mathsf{M}_0$–frame)
A quasi–$\mathbf{IL}\mathsf{M}_0$–frame is a quasi–frame that satisfies the following additional properties.

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

LEMMA 6.11
If F is a quasi–**IL**M_0–frame, then $K = (K^1)^{tr}$.

PROOF. Using Lemma 6.8. ∎

LEMMA 6.12
Suppose that F is a quasi–**IL**M_0–frame. Let $K = K(F)$. Let K', K'' and K''' the smallest binary relations on W satifying 1. and 2. of 6.6 and additionaly we have the following.

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

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

PROOF. Using Lemma 6.11. ∎

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

COROLLARY 6.13
If $F = \langle W, R, S, \nu\rangle$ is an adequate **IL**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–**IL**M_0–frame

LEMMA 6.14 (**IL**M_0–closure)
Any quasi–**IL**M_0–frame can be extended to an adequate **IL**M_0–frame.

PROOF. Given a quasi–**IL**M_0–frame F we construct a sequence

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

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

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

In this case we set, of course, $F_{n+1} := \langle W_n, R_n \cup \langle a, c\rangle, S_n, \nu_n\rangle$. First we will show by induction that each F_n is a quasi–**IL**M_0–frame. Then we show that the union $\hat{F} = \bigcup_{n\geq 0} F_n$, is quasi and satisfies all the **IL**M_0 frame conditions.

We assume that F_n is a quasi-**IL**M_0-frame and define $K^n := K(F_n)$, $R^n := R^{F_n}$ and $S^n := S^{F_n}$. Quasi-ness of F_{n+1} will follow from Claim 6.15, and from Claim 6.16 we may conlude that F_{n+1} is indeed a quasi-**IL**M_0-frame.

Claim 6.15
For all w, x, y and A we have the following.

(a) $R^{n+1} \subseteq K^n$
(b) $x(S_w^{n+1} \cup R^{n+1})^* y \Rightarrow x(S_w^n \cup K^n)^* y$
(c) $F_{n+1} \models x \in \mathcal{C}_w^A \Rightarrow F_n \models x \in \mathcal{N}_w^A$.

Proof. We distinguish cases according to which imperfection is dealt with in the step from F_n to F_{n+1}. The only interesting case is the new imperfection which is dealt with by Lemma 6.12, Item $3''$. ∎

Claim 6.16
For all w, x and A we have the following.

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

Proof. Item 1. follows by Claim 6.15 and Lemma 6.7. Item 2. follows from Item 1. and Claim 6.15-(b). Item 3. is an immediate corollary of item 2. ∎

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

Lemma 6.17
Let $F = \langle W, R, S, \nu \rangle$ be a quasi–$\mathbf{IL}\mathsf{M}_0$–frame and $K = K(F)$. Then

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

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

It is easy to see that K' satisfies properties 1., 2. and 3. of Definition 6.6; It follows on the two observations on T we just made. Since K is the smallest binary relation that satisfies these properties we conclude $K \subseteq K'$. ∎

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 6.18
Let $F = \langle W, R, S, \nu \rangle$ be a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. And let $G = \langle W^-, R^-, S^-, \nu^- \rangle$ be a subframe of F (which means $W^- \subseteq W$, $R^- \subseteq R$, $S^- \subseteq S$ and $\nu^- \subseteq \nu$). If

(a) for each $f \in W - W^-$ and $g \in W^-$ not $f(R \cup S)g$ and
(b) $R{\restriction}_{W^-} \subseteq K(G)$.

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

Proof. Clearly $K(F){\restriction}_{W^-}$ satisfies the properties 1., 2. and 3. of the definition of $K(G)$ (Definition 6.6). Thus, since K_G is the smallest such relation, we get that $K(G) \subseteq K(F){\restriction}_{W^-}$.

Let $K' = K(F) - (K(F){\restriction}_{W^-} - K(G))$. Using Lemma 6.17 one can show that $K(F) \subseteq K'$. From this it immediately follows that $K(F){\restriction}_{W^-} \subseteq K(G)$. ∎

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

LEMMA 6.19
Let Γ and Δ be MCS's. $\Gamma \prec_C \Delta$.

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

There exist $k \leq n$. 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 (i.e. $\Delta \subseteq_\Box \Delta_i$),
- $Q \in \Delta_0$,
- 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

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

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

$$Q, \overline{S}_I, \Box A, \neg B \wedge \Box\neg B \vdash \bot \tag{6.1}$$

or,

$$\text{for some } i \notin I, \quad T_i, \overline{S}_I, \Box A, \neg B \wedge \Box\neg B \vdash \bot. \tag{6.2}$$

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

$$T_{i_k}, \overline{S}_{I_k}, \Box A, \neg B \wedge \Box\neg B \vdash \bot. \tag{6.3}$$

Additionally, we will verify that for each k

$$\text{(6.1) does not hold with } I_k \text{ for } I.$$

We will define i_k with induction on k. We define $I_1 = \emptyset$. And by Lemma 4.18, (6.1) does not hold with $I = \emptyset$. Moreover, because of this, (6.2) must be true with $I = \emptyset$. So, there exists some $i \in \{1, \ldots, n\}$ such that

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

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

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

$$Q, \overline{S}_{I_{k+1}}, \Box A, \neg B \wedge \Box\neg B \nvdash \bot. \tag{6.4}$$

Let us suppose that this is not so. Then

$$\vdash \Box(Q \to \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_k}). \tag{6.5}$$

So,

$$\begin{aligned}
\Gamma \vdash P &\rhd Q \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee S_{i_k} && \text{by (6.5)} \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee T_{i_k} \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} \vee (T_{i_k} \wedge \Box A \wedge \neg B \wedge \Box\neg B \wedge \overline{S}_{I_k}) \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \vee \cdots \vee S_{i_{k-1}} && \text{by (6.3)} \\
&\quad \vdots \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee S_{i_1} \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee T_{i_1} \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B \vee (T_{i_1} \wedge \Box A \wedge \neg B \wedge \Box\neg B) \\
&\rhd \Diamond\neg A \vee B \vee \Diamond B. && \text{by (6.3), with } k = 1.
\end{aligned}$$

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 4.18 there exists some MCS Δ with $\Gamma \prec_C \Delta$ that contains $B \vee \Diamond B$. This is a contradiction, so we have shown (6.4).

But now, since (6.4) is indeed true, and thus (6.1) with I_{k+1} for I is false, (6.2) 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 $I = \{1, 2, \ldots, n\}$, (6.2) cannot be true. Thus, for $I = \{1, 2, \ldots, n\}$, (6.1) 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. ∎

In order to formulate the invariants needed in the main lemma applied to $\mathbf{IL}\mathsf{M}_0$, we need one more definition and a corollary.

Definition 6.20 ($\subset_1$, $\subset$)
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 6.21
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 $x \subset y \Rightarrow \nu(x) \subseteq_{\Box} \nu(y)$.

6.3 Frame condition

The following theorem is well known.

THEOREM 6.22
For an **IL**-frame $F = \langle W, R, S, \nu \rangle$ we have

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

6.4 Invariants

Let $\mathcal{D}$ be some finite set of formulas, closed under subformulas and single negation.
During the construction we will keep track of the following main–invariants.

$\mathcal{I}_{\Box}$ for all y, $\{\nu(x) \mid xK^1y\}$ is linearly ordered by $\subseteq_{\Box}$
$\mathcal{I}_{\mathrm{d}}$ $wK^1x \wedge wK^{\geq 2}x'(S_w \cup K)^*x \to$ 'there does not exists a deficiency in w w.r.t. x'
$\mathcal{I}_S$ $wKxKy(S_w \cup K)^*y' \to$
'the $\subseteq_{\Box}$-max of $\{\nu(t) \mid wKtK^1y'\}$, if it exists, is $\subseteq_{\Box}$-larger than $\nu(x)$'
$\mathcal{I}_{\mathcal{N}}$ $wKxKy \wedge y \in \mathcal{N}^A_w \to x \in \mathcal{N}^A_w$
$\mathcal{I}_{\mathcal{D}}$ $xRy \to \exists\, A \in (\nu(y) \setminus \nu(x)) \cap \{\Box D \mid D \in \mathcal{D}\}$
$\mathcal{I}_{\mathsf{M}_0}$ All conditions for an adequate $\mathbf{IL}\mathsf{M}_0$–frame hold

In order to ensure that the main–invariants are preserved during the construction we need to consider the following sub–invariants.[13]

$\mathcal{J}_{\mathrm{u}}$ $wK^{\geq 2}x(\mathcal{S}_w)^{\mathrm{tr}}y \wedge wK^{\geq 2}x'(\mathcal{S}_w)^{\mathrm{tr}}y \to x = x'$
$\mathcal{J}_{K^1}$ $wKxK^1y(\mathcal{S}_w)^{\mathrm{tr}}y'K^1z \to xK^1z$
$\mathcal{J}_{\subset}$ $y \subset x \wedge x \subset y \to y = x$
$\mathcal{J}_{\mathcal{N}_1}$ $x(\mathcal{S}_v)^{\mathrm{tr}}y \wedge wKy \wedge x \in \mathcal{N}^A_w \to y \in \mathcal{N}^A_w$
$\mathcal{J}_{\mathcal{N}_2}$ $x(\mathcal{S}_w)^{\mathrm{tr}}y \wedge y \in \mathcal{N}^A_w \to x \in \mathcal{N}^A_w$
$\mathcal{J}_{\nu_1}$ '$\nu(w,y)$ is defined' $\wedge\ vKy \to v \subset w$
$\mathcal{J}_{\nu_2}$ '$\nu(w,y)$ is defined' $\to wK^1y$
$\mathcal{J}_{\nu_4}$ If $x(\mathcal{S}_w)^{\mathrm{tr}}y$, 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 Section 6.1. $\mathcal{I}_{\mathsf{M}_0}$ is there to ensure that our final frame is an $\mathbf{IL}\mathsf{M}_0$–frame. About the

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

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 6.23
Any quasi–$\mathbf{IL}\mathsf{M}_0$–frame that satisfies all of the above invariants, except possibly $\mathcal{I}_{\mathsf{M}_0}$, can be extended to an $\mathbf{IL}\mathsf{M}_0$–frame that satisfies all the invariants.

Proof. Only $\mathcal{I}_{\mathcal{D}}$ and $\mathcal{I}_{\mathrm{d}}$ need some attention. All the other invariants are given in terms of relations that do not change during the construction of the $\mathbf{IL}\mathsf{M}_0$-closure (Lemma 6.14). ∎

Lemma 6.24
Let $F = \langle W, R, S, \nu\rangle$ be a quasi–$\mathbf{IL}\mathsf{M}_0$–frame. Then $F \models x \in \mathcal{N}^A_w$ iff. one of the following cases applies.

1. $\nu(w, x) = A$
2. There exists $t \in \mathcal{N}^A_w$ such that tKx
3. There exists $t \in \mathcal{N}^A_w$ such that $t\mathcal{S}_w x$

Corollary 6.25
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 $x \in \mathcal{N}^A_w$ implies $\nu(w, x) = A$ or there exists some $t \in \mathcal{N}^A_w$ such that tKx.

Lemma 6.26
Let F be a quasi–frame which satisfies $\mathcal{J}_{\mathcal{N}_2}$, $\mathcal{J}_{\nu_1}$, $\mathcal{J}_{\nu_3}$ and $\mathcal{J}_{\nu_4}$. Then $x\mathcal{S}_v y$, $y \in \mathcal{N}^A_w \Rightarrow x \in \mathcal{N}^A_w$.

Proof. Suppose $x\mathcal{S}_v y$ and $y \in \mathcal{N}^A_w$. Then, by Corollary 6.25, $\nu(w, y) = A$ or, for some $t \in \mathcal{N}^A_w$, tKy. In the first case we obtain $w = v$ by $\mathcal{J}_{\nu_3}$ and $\mathcal{J}_{\nu_4}$. And thus by $\mathcal{J}_{\mathcal{N}_2}$, $x \in \mathcal{N}^A_w$. In the second case we have, by $\mathcal{J}_{\nu_4}$ and $\mathcal{J}_{\nu_1}$ that $t \subset v$. Which implies, by Lemma 6.21–1., tKx. ∎

6.5 Solving problems

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

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

and define $\hat{K} = K(\hat{F})$. The frames F and $\hat{F}$ satisfy the conditions of Lemma 6.18. Thus we have

$$\forall xy{\in}F\ xKy \Leftrightarrow x\hat{K}y. \tag{6.6}$$

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

Claim 6.27
$\hat{F}$ is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame.

Proof. A simple check of Properties (1.–5.) of Definition 5.1 (quasi–frames) and Properties (6.–10.) of Definition 6.10 (quasi–$\mathbf{IL}\mathsf{M}_0$–frames) and the remaining ones in Definition 5.1 (quasi–frames). Let us comment on two of them.

$x\hat{K}y \to \hat{\nu}(x) \prec \hat{\nu}(y)$ follows from Lemma 6.17 and (6.6).

Let us show $\hat{F} \models x \in \mathcal{N}^C_w \Rightarrow \hat{\nu}(w) \prec_C \hat{\nu}(x)$. We have $\forall xw{\in}F\ F \models x \in \mathcal{N}^C_w \Leftrightarrow \hat{F} \models x \in \mathcal{N}^C_w$. So we only have to consider the case $\hat{F} \models \mathbf{b} \in \mathcal{N}^C_w$. If $w = \mathbf{a}$ then we are done by choice of $\hat{\nu}(\mathbf{b})$. Otherwise, by Lemma 6.26, we have for some $x \in F$, $F \models x \in \mathcal{N}^C_w$ and $x\hat{K}\mathbf{b}$. By the first property we proved, we get $\hat{\nu}(x) \prec \hat{\nu}(\mathbf{b})$. So, since $\hat{\nu}(w) \prec_C \hat{\nu}(x)$ we have $\hat{\nu}(w) \prec_C \hat{\nu}(\mathbf{b})$. ∎

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

Lemma 6.28
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 $wK^{\geq 2}u(\mathcal{S}_w)^{\mathrm{tr}}\mathbf{a}$.

Proof. If such w and u do not exists then $T = K \cup \{\mathbf{a}, \mathbf{b}\}$ satisfies the conditions of Lemma 6.8. In which case $xK^1\mathbf{b}$ gives $xT\mathbf{b}$ which implies $x = \mathbf{a}$. The uniqueness of w follows from $\mathcal{J}_{\nu_3}$ and $\mathcal{J}_{\nu_4}$. The uniqueness of u follows from $\mathcal{J}_{\mathrm{u}}$ and the uniqueness of w. ∎

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

Lemma 6.29
For all x. If $x\hat{K}^1\mathbf{b}$ then $x \subset \mathbf{a}$.

Proof. Let $K' = K \cup \{(x, \mathbf{b}) \mid x\hat{K}\mathbf{b} \wedge x \subset \mathbf{a}\}$. It is not hard to show that K' satisfies the conditions of T in Lemma 6.8. ∎

Lemma 6.30
Suppose the conditions of Lemma 6.28 are satisfied and let $\mathbf{u}$ be the u asserted to exist. Then for all $x \neq \mathbf{a}$, if $x\hat{K}^1\mathbf{b}$, then $xK^1\mathbf{u}$.

Proof. By Lemma 6.29 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 6.21, $xKx_{i+1}K\mathbf{b}$. So, $xK^{\geq 2}\mathbf{b}$. A contradiction. The lemma now follows by showing, with induction on i and using $F \models \mathcal{J}_{K^1}$, that for all $i \geq 0$, $x_{n-(i+1)}K^1\mathbf{u}$.
∎

Lemma 6.31
$\hat{F}$ satisfies all the sub-invariants.

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

$\mathcal{J}_{\nu_1}$ follows from Lemma 6.29, so let us treat $\mathcal{J}_{K^1}$. Suppose $w\hat{K}x\hat{K}^1y(\hat{\mathcal{S}}_w)^{\mathrm{tr}}y'\hat{K}^1z$. We can assume that at least one of w, x, y, y', z is not in F and the only candidate for this is z. So we have $z = \mathbf{b}$. We can assume that $x \neq y'$ (otherwise we are done at once), so the conditions of Lemma 6.28 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 6.30, $t = \mathbf{a}$ 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 6.21 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 6.30 and since $y'\hat{K}^1 z$ we have $y'\hat{K}^1\mathbf{u}$ or $y' = \mathbf{a}$. In the first case, since $F \models \mathcal{J}_{K^1}$, we have $x\hat{K}^1\mathbf{u}$. In the second case we obtain, by the uniqueness of $\mathbf{u}$, that $y = \mathbf{u}$ and thus $x\hat{K}^1\mathbf{u}$. So in any case we have

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

A contradiction.

■

Lemma 6.32
Possibly with the exception of $\mathcal{I}_{\mathsf{M}_0}$, $\hat{F}$ satisfies all the main-invariants.

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

First we treat $\mathcal{I}_\Box$. So we have to show that for all y, $\{\hat{\nu}(x) \mid x\hat{K}^1 y\}$ is linearly ordered by $\subseteq_\Box$. We only need to consider the case $y = \mathbf{b}$. If $\{\mathbf{a}\} = \{x \mid x\hat{K}^1\mathbf{b}\}$ then the claim is obvious. So we can assume that the condition of Lemma 6.28 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 6.17, $x\hat{K}^1\mathbf{b} \Rightarrow x \subseteq_\Box \mathbf{a}$.

Now we look at $\mathcal{I}_\mathcal{N}$: $w\hat{K}x\hat{K}y \wedge \hat{F} \models y \in \mathcal{N}_w^A \rightarrow \hat{F} \models x \in \mathcal{N}_w^A$. 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 6.24, $\hat{\nu}(w, \mathbf{b}) = A$ 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

$$t\hat{K}^1\mathbf{b}. \tag{6.7}$$

We suppose that the conditions of Lemma 6.28 are fulfilled (the other case is easy). If $t\hat{K}^1\mathbf{u}$ and $x\hat{K}^*\mathbf{u}$ then we are done simmilarly as the case above. So assume $t\hat{K}^1\mathbf{a}$ or $x\hat{K}^*\mathbf{a}$. Since wRt and wRx in any case we have $w\hat{K}\mathbf{a}$. Now by Lemma 6.26 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 (6.7), $\mathbf{u} \in \mathcal{N}_w^A \vee \mathbf{a} \in \mathcal{N}_w^A$. 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 $F \models \mathcal{I}_\mathcal{N}$.

■

To finish this subsection we note that by Lemma 6.14 and Corollary 6.23 we can extend $\hat{F}$ to an adequate $\mathbf{IL}\mathsf{M}_0$–frame that satisfies all invariants.

6.6 Solving deficiencies

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

Suppose $\mathbf{a}R^{\geq 2}\mathbf{b}$ (the case $\mathbf{a}R^1\mathbf{b}$ is easy). Let x be the $\subseteq_\Box$-maximum of $\{x \mid \mathbf{a}KxK^1\mathbf{b}\}$. This maximum exists by $\mathcal{I}_\Box$. Pick some A such that $\mathbf{b} \in \mathcal{N}_\mathbf{a}^A$. (If such an A exists, then by adequacy of F, it is unique. If no such A exists, take $A = \bot$.) By $\mathcal{I}_\mathcal{N}$ and adequacy we have $\nu(\mathbf{a}) \prec_A \nu(x)$. So we have $C \rhd D \in \nu(\mathbf{a}) \prec_A \nu(x) \ni \Diamond C$.

We apply Lemma 6.19 to obtain, for some set Y, disjoint from W, a set $\{\Delta_y \mid y \in Y\}$ of MCS's with all the properties as stated in that lemma. We define

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

Claim 6.33
$\hat{F}$ is a quasi–$\mathbf{IL}\mathsf{M}_0$–frame.

Proof. An easy check of Properties (1.–5.) of Definition 5.1 (quasi–frames) and Properties (6.–10.) of Definition 6.10 (quasi–$\mathbf{IL}\mathsf{M}_0$–frames). Let us comment on two cases.

First we see that $x\hat{K}y \to \hat{\nu}(x) \prec \hat{\nu}(y)$. We can assume $y \in Y$. By Lemma 6.17 we obtain some z with $\hat{\nu}(x) \subseteq_\Box \hat{\nu}(z)$ and $x(\hat{R} \cup \hat{S})^* z\hat{R}y$. This z can only be $\mathbf{a}$. By choice of $\hat{\nu}(y)$ we have $\hat{\nu}(\mathbf{a}) \prec \hat{\nu}(y)$. And thus $\hat{\nu}(x) \prec \hat{\nu}(y)$.

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

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

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

Lemma 6.34
For any $x \in \hat{F}$ and $y \in Y$ we have $x\hat{K}^1y \to x \subset \mathbf{a}$.

Proof. We put $K' = K \cup \{(x, y) \mid y \in Y,\ x\hat{K}y,\ x \subset \mathbf{a}\}$. By showing that K' satisfies the conditions of T in Lemma 6.8. we obtain $x\hat{K}^1y \to xK'y$. So if $x\hat{K}^1y$ then $xK'y$. But if $y \in Y$ then xKy does not hold. Thus we have $x \subset \mathbf{a}$. ∎

Lemma 6.35
Suppose $y \in Y$ and $\mathbf{a}\hat{K}^1z$. Then for all x, $x\hat{K}^1y \to x\hat{K}^1z$.

Proof. Suppose xK^1y. By Lemma 6.34 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_iKx_{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 $i \leq n$, $x_{n-i}K^1z$. ∎

Lemma 6.36
$\hat{F}$ satisfies all the sub-invariants.

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

First we show $\mathcal{J}_{\mathrm{u}}$: $w\hat{K}^{\geq 2}x(\hat{\mathcal{S}}_w)^{tr}y \wedge w\hat{K}^{\geq 2}x'(\hat{\mathcal{S}}_w)^{tr}y \to x = x'$. Suppose that $w\hat{K}^{\geq 2}x(\hat{\mathcal{S}}_w)^{tr}y$ and $w\hat{K}^{\geq 2}x'(\hat{\mathcal{S}}_w)^{tr}y$. We can assume that $y \in Y$. (Otherwise all of w, x, x', y are in F and we are done by $F \models \mathcal{J}_{\mathrm{u}}$.) We clearly have $w \in F$. If $x \in Y$ then $w = \mathbf{a}$ and thus $w\hat{K}^1x$. 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 $wK^{\geq 2}x'(\mathcal{S}_w)^{tr}\mathbf{b}$. By $F \models \mathcal{J}_{\nu_4}$ we now obtain that $\nu(w, \mathbf{b})$ is defined. And thus by $F \models \mathcal{J}_{\nu_2}$, $wK^1\mathbf{b}$. A contradiction.

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

Now let us see that $\mathcal{J}_{K^1}$ holds, that is $w\hat{K}x\hat{K}^1y(\hat{\mathcal{S}}_w)^{tr}y'\hat{K}^1z \to x\hat{K}^1z$. Suppose $w\hat{K}x\hat{K}^1y(\hat{\mathcal{S}}_w)^{tr}y'\hat{K}^1z$. 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 6.35 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. ∎

Lemma 6.37
Except for $\mathcal{I}_{\mathsf{M}_0}$, $\hat{F}$ satisfies all main-invariants.

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

First we show $\mathcal{I}_\Box$: For all y, $\{\hat{\nu}(x) \mid x\hat{K}^1y\}$ is linearly ordered by $\subseteq_\Box$. Let $y \in \hat{F}$ and consider the set $\{x \mid xK^1y\}$. Since $\hat{K}\restriction_F = K$ and for all $y \in Y$ there does not exists z with $y\hat{K}^1z$ we only have to consider the case $y \in Y$. Fix some a_1 such that $\mathbf{a}K^1a_1K^*\mathbf{b}$. By Lemma 6.34 for any such y we have

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

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

Now let us see $\mathcal{I}_\mathcal{N}$: $w\hat{K}x\hat{K}y \wedge \hat{F} \models y \in \mathcal{N}^A_w \to \hat{F} \models x \in \mathcal{N}^A_w$. 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 6.34, $x \subset \mathbf{a}$. So, $wKxK\mathbf{b}$. By Lemma 6.26, $F \models \mathbf{b} \in \mathcal{N}^A_w$ and thus $\hat{F} \models x \in \mathcal{N}^A_w$. ∎

To finish this section we noting that by Lemma 6.14 and Corollary 6.23 we can extend $\hat{F}$ to an adequate $\mathbf{IL}\mathsf{M}_0$–frame that satisfies all invariants.

6.7 Rounding up

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

7 The logic ILW*

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

Theorem 7.1
$\mathbf{IL}\mathsf{W}^*$ is a complete logic.

For a long time $\mathbf{IL}\mathsf{W}^*$ has been conjectured ([29]) to be $\mathbf{IL}$(All). A first step in proving this conjecture would have been a modal completeness result. However, the modal completeness of $\mathbf{IL}\mathsf{W}^*$ resisted many attempts as the modal completeness of $\mathbf{IL}\mathsf{M}_0$, which is an essential part of $\mathbf{IL}\mathsf{W}^*$, was so hard and involved. (In [9] a completeness proof for $\mathbf{IL}\mathsf{W}$ was given.)

Finally, now that all the machinery has been developed, a modal completeness proof for $\mathbf{IL}\mathsf{W}^*$ can be given. The completeness proof of $\mathbf{IL}\mathsf{W}^*$ lifts almost completely along with the completeness proof for $\mathbf{IL}\mathsf{M}_0$. We only need some minor adaptations.

7.1 Preliminaries

The frame condition of W is well known.

THEOREM 7.2
For any **IL**-frame F we have that $F \models \mathsf{W} \Leftrightarrow \forall w\ (S_w; R)$ is conversely well-founded.

We can define a new principle M_0^* that is equivalent to W^*, 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 7.3
$\mathbf{IL}\mathsf{M}_0\mathsf{W} = \mathbf{IL}\mathsf{W}^* = \mathbf{IL}\mathsf{M}_0^*$

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

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

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. $B \wedge \Box C \rhd (B \wedge \Box C \wedge \Box\neg A) \vee (\Box C \wedge \Diamond A) \rhd B \wedge \Box C \wedge \Box\neg A$. ∎

COROLLARY 7.4
For any **IL**-frame we have that $F \models \mathsf{W}^*$ iff. both (for each w, $(S_w; R)$ is conversely well-founded) and $(\forall w, x, y, y', z\ (wRxRyS_wy'Rz \to xRz))$.

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

DEFINITION 7.5 ($\subsetneq_\Box^{\mathcal{D}}$)
Let $\mathcal{D}$ be a finite set of formulas. Let $\subsetneq_\Box^{\mathcal{D}}$ be a binary relation on MCS's defined as follows. $\Delta \subsetneq_\Box^{\mathcal{D}} \Delta'$ iff.

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

LEMMA 7.6
Let F be a quasi-frame and $\mathcal{D}$ be a finite set of formulas. If $wRxRyS_wy' \to \nu(x) \subsetneq_\Box^{\mathcal{D}} \nu(y')$ then $(R; S_w)$ is conversely well-founded.

PROOF. By the finiteness of $\mathcal{D}$. ∎

Lemma 7.7
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 7.8 (Adequate $\mathbf{IL}\mathsf{W}^*$-frame)
Let $\mathcal{D}$ be a set of formulas. We say that an adequate $\mathbf{IL}\mathsf{M}_0$-frame is an adequate $\mathbf{IL}\mathsf{W}^*$-frame (w.r.t. $\mathcal{D}$) iff. the following additional property holds.

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

Definition 7.9 (Quasi-$\mathbf{IL}\mathsf{W}^*$-frame)
Let $\mathcal{D}$ be a set of 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 holds.

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.

Lemma 7.10
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 6.14. 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. Additionally we have to show that $\hat{F}$ is an adequate $\mathbf{IL}\mathsf{W}^*$-frame.

But this is rather trivial. As noted in the proof of Lemma 6.14, The relation K and the relations $(\mathcal{S}_w)^{tr}$ 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 bounded chains. So, $\hat{F}$ is an adequate $\mathbf{IL}\mathsf{W}^*$-frame. ∎

Lemma 7.11
Let Γ and Δ be MCS's with $\Gamma \prec_C \Delta$,

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

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

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

Proof. The proof is a straightforward adaptation of the proof of Lemma 6.19. In that proof, a trick was to postpone an application of M_0 as long as possible. We do the same here but let an application of M_0 on $P \rhd \Diamond P \vee \psi$ be preceded by an application of W to obtain $P \rhd \psi$. ∎

7.2 Completeness

Again, we specify the four ingredients from Remark 4.19. The **Frame condition** is contained in Corollary 7.4.

The **Invariants** are all those of $\mathbf{IL}\mathsf{M}_0$ and additionally

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

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

Problems. We have to show that we can solve problems in an adequate $\mathbf{IL}\mathsf{W}^*$-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. (This is basically Lemma 7.10.)

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 7.11 instead of Lemma 6.19.

Rounding up. We have to show that the union of a bounded chain of quasi-$\mathbf{IL}\mathsf{W}^*$-frames that satisfy all the invariants is an $\mathbf{IL}\mathsf{W}^*$-frame. The only novelty is that we have to show that in this union for each w we have that $(R; S_w)$ is conversely well-founded. But this is ensured by $\mathcal{I}_{w^*}$ and Lemma 7.6.

References

[1] C. Areces, D. de Jongh, and E. Hoogland. The interpolation theorem for IL and ILP. In *Proceedings of AiML98. Advances in Modal Logic*, Uppsala. Sweden, October 1998. Uppsala University.

[2] L.D. Beklemishev. On the complexity of arithmetic interpretations of modal formulae. *Archive for Mathematical Logic*, 32:229–238, 1993.

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

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

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

[6] 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.

[7] D. de Jongh and D. Pianigiani. Solution of a problem of David Guaspari. *Studia Logica*, 1996.

[8] D.H.J. de Jongh and F. Veltman. Provability logics for relative interpretability. In *[21]*, pages 31–42, 1990.

[9] D.H.J. de Jongh and F. Veltman. Modal completeness of ILW. In J. Gerbrandy, M. Marx, M. Rijke, and Y. Venema, editors, *Essays dedicated to Johan van Benthem on the occasion of his 50th birthday.* Amsterdam University Press, Amsterdam, 1999.

[10] D.H.J. de Jongh and A. Visser. Explicit fixed points in interpretability logic. *Studia Logica*, 50:39–50, 1991.

[11] E. Goris. Extending ILM with an operator for Σ_1–ness. Illc prepublication series, University of Amsterdam, 2003.

[12] D. Guaspari. Sentences implying their own provability. *Journal of Symbolic Logic*, 48:777–789, 1983.

[13] P. Hájek and P. Pudlák. *Metamathematics of First Order Arithmetic.* Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[14] R. Hirsch and I. Hodkinson. *Relation Algebras by Games*, volume 147 of *Studies in Logic.* Elsevier, North-Holland, 2002.

[15] I. Hodkinson, S. Mikulás, and Y. Venema. Axiomatizing complex algebras by games. *Algebra Universalis*, 46:455–478, 2001.

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

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

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

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

[20] 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.

[21] P.P. Petkov, editor. *Mathematical logic, Proceedings of the Heyting 1988 summer school in Varna, Bulgaria*. Plenum Press, Boston, 1990.

[22] P. Pudlák. Cuts, consistency statements and interpretations. *Journal of Symbolic Logic*, 50:423–441, 1985.

[23] V. Shavrukov. The logic of relative interpretability over Peano arithmetic (in Russian). Technical Report Report No.5, Steklov Mathematical Institute, Moscow, 1988.

[24] R.M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 28:33–71, 1976.

[25] A. Tarski, A. Mostowski, and R. Robinson. *Undecidable theories*. North–Holland, Amsterdam, 1953.

[26] V. Švejdar. Some independence results in interpretability logic. *Studia Logica*, 50:29–38, 1991.

[27] A. Visser. Preliminary notes on interpretability logic. Technical Report LGPS 29, Department of Philosophy, Utrecht University, 1988.

[28] A. Visser. Interpretability logic. In *[21]*, pages 175–209, 1990.

[29] A. Visser. The formalization of interpretability. *Studia Logica*, 50(1):81–106, 1991.

[30] A. Visser. A course on bimodal provability logic. *Annals of Pure and Applied Logic*, pages 109–142, 1995.

[31] A. Visser. An overview of interpretability logic. In M. Kracht, M. de Rijke, and H. Wansing, editors, *Advances in modal logic '96*, pages 307–359. CSLI Publications, Stanford, CA, 1997.

[32] 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.

[33] M Vuković. Some correspondence of principles in interpretability logic. *Glasink Matematički*, 31(51):193–200, 1996.

Received 23-1-2006