MODULAR SEQUENT CALCULI FOR INTERPRETABILITY LOGICS

COSIMO PERINI BROGI
SySMA Unit, IMT School for Advanced Studies Lucca

SARA NEGRI
Department of Mathematics, University of Genoa
and

NICOLA OLIVETTI
LIS, Aix-Marseille University, CNRS

Abstract. An original family of labelled sequent calculi $\mathsf{G3IL}^\star$ for classical interpretability logics is presented, modularly designed on the basis of Verbrugge semantics (a.k.a. generalised Veltman semantics) for those logics. We prove that each of our calculi enjoys excellent structural properties, namely, admissibility of weakening, contraction and, more relevantly, cut. A complexity measure of the cut is defined by extending the notion of range previously introduced by Negri w.r.t. a labelled sequent calculus for Gödel–Löb provability logic, and a cut-elimination algorithm is discussed in detail. To our knowledge, this is the most extensive and structurally well-behaving class of analytic proof systems for modal logics of interpretability currently available in the literature.

§1. Introduction.

1.1. Background. Interpretations arise in several (meta)mathematics areas, and many variations exist.[1] For instance, it is possible to interpret propositional intuitionistic logic into classical Gödel–Löb logic GL by establishing an equivalence between the axiomatisation $\mathbb{GL}$ and the axiomatic calculus $\mathbb{IPC}$; from that, one could also interpret $\mathbb{IPC}$ into an arithmetical theory T that is adequate to $\mathbb{GL}$.

An even simpler example is given by Gödel's numbering, which, according to Tarski [47], interprets (a model for) meta-mathematical reasoning into the standard model for arithmetic by defining an injective function that maps finite strings of arithmetical symbols into $\mathbb{N}$, and a further function mapping each meta-predicate into its arithmetical counterpart.

Received: January 27, 2024.
2020 *Mathematics Subject Classification*: Primary 03F05, 03F07, 03F45, Secondary 03B45, 03F25.
Key words and phrases: interpretability logics, structural proof theory, cut-elimination, Verbrugge semantics, modal logics for metamathematics.

[1] A methodological treatment of the notion was presented first in [48], where the basic properties of this concept are introduced.

doi:10.1017/S1755020325100701

In the present work, we will assume a very general version of that concept: An *interpretation* of a theory T into a theory T' is just a structure-preserving translation t such that if $\mathsf{T} \vdash A$ then $\mathsf{T}' \vdash t(A)$. More precisely, we will consider interpretability between arithmetical theories of the form $\mathsf{T} + A$, where T is some base theory satisfying the standard Hilbert–Bernays–Löb provability conditions as discussed in e.g., [5, 42, 43], and A is a schema in the language of T.[2]

Modal logics for interpretability arise as an extension of the language of provability logic through a binary modal operator $\rhd$ capturing the relation of (relative) interpretability between two arithmetical theories.

Recall first that, in provability logic, we denote by $*$ a standard arithmetical realisation of the modal language, as described in e.g., [51]: In particular, $*$ commutes with the classical operators and maps $\Box A$ into $Bew(\ulcorner A^* \urcorner)$, where $Bew(x)$ is the standard provability predicate for the given arithmetical theory T.

In interpretability logics, the propositional formula $A \rhd B$ is then intended as the modal counterpart of the arithmetical formula $Int_{\mathsf{T}}(\ulcorner A^* \urcorner, \ulcorner B^* \urcorner)$—where $Int_{\mathsf{T}}(x, y)$ is the formal predicate in the language of T for relative interpretability over T itself—expressing the fact that the arithmetical theory T extended by A^* interprets the arithmetical theory T extended by B^*.

The origins of interpretability logics date back to [54], which axiomatised the basic modal framework by extending the language of $\mathbb{GL}$ with some minimal axioms for the core logic for interpretability $\mathbb{IL}$, on top of which several systems can be constructed.

Completeness results w.r.t. a Kripke-style relational semantics, based on so-called Veltman frames, were presented first in [7] for $\mathbb{IL}$ and some extensions by using a canonical model construction. More complex proofs were developed by the same authors in 1999, and further techniques were introduced to achieve subsequent completeness results since the beginning of the 2000s starting with [12]. Recent results have been obtained also for subsystems of $\mathbb{IL}$ in [16, 19, 34].

On the arithmetical side, by tweaking the proof strategy in [45] for $\mathbb{GL}$, it is possible to prove also the *arithmetical* completeness of some extensions of $\mathbb{IL}$. The interpretability logic for T is, as expected,

$$\mathsf{IL}(\mathsf{T}) := \{A \mid \text{for any realisation } *, \mathsf{T} \vdash A^*\}.$$

In [41], it is proven that this notion is Σ^0_3-complete. It is also known, that $\mathbb{IL}$ is not the interpretability logic for any arithmetical theory T. *However*, if we assume that T is Σ_1-sound and proves full induction, by adding the schema

$$\mathsf{M} := A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C,$$ [3]

we have that

$$\mathsf{IL}(\mathsf{T}) = \mathbb{ILM} := \mathbb{IL} + \mathsf{M}.$$

[2] This is usually called *relative interpretability*. Notice that, according to the definition, provability is a special case of relative interpretability: $\mathsf{T} \vdash A$ if and only if $\mathsf{T} + \bot$ is interpretable in $\mathsf{T} + \neg A$.

[3] In parsing formulas, we assume that $\rhd$ binds stronger than $\to$, but weaker than the other connectives.

Similarly if we add to $\mathbb{IL}$ the schema

$$\mathsf{P} := A \rhd B \to \Box(A \rhd B),$$

then

$$\mathbb{IL} + \mathsf{P} =: \mathbb{ILP} = \mathsf{IL}(\mathsf{T})$$

for any T that is Σ_1-sound, finitely axiomatised, and such that it proves the totality of `supexp`.[4]

The most intriguing aspect of interpretability logics is this sensitivity to the base arithmetical theory. The main open question in the field is indeed establishing the interpretability logic of all reasonable arithmetical theories, i.e.,[5]

$$\mathsf{IL}(All) := \{A \mid \forall \mathsf{T} \supseteq \mathsf{I}\Delta_0 + \mathsf{exp}, \text{ for any realisation } *, \mathsf{T} \vdash A^*\}.$$

What we know, after [55], is that $\mathbb{IL} \subset \mathsf{IL}(All) \subset \mathbb{ILM} \cap \mathbb{ILP}$, but a modal characterisation of $\mathsf{IL}(All)$ is still unknown.

1.2. This article. There are many further open questions in the field of interpretability logics. In our setting, it is worth noticing that very little is known about the structural proof theory for interpretability logics: Sasaki [39] gives a Gentzen-style sequent calculus for $\mathbb{IL}$ only, while Hakoniemi and Joosten [15] presents a labelled tableaux system for some extensions of $\mathbb{IL}$ based on standard Veltman semantics.[6]

In the present paper, we extend the proof-theoretic analysis of interpretability logics by introducing a family $\mathsf{G3IL}^\star$ of labelled sequent calculi which cover a wide range of modal systems for interpretability naturally. Their design is based on the methodology of Negri [28, 31] but, instead of working with formal relational semantics or formal neighbourhood semantics, these original calculi internalise the hybrid models by Rineke Verbrugge, also called generalised Veltman structures in [50].

One of the main advantages of reasoning on interpretability logics by using Verbrugge semantics instead of standard Veltman semantics is that the former subsumes the latter, but it is also capable of distinguishing interpretability principles that are equivalent in a model theory based on Veltman frames. Moreover, some interpretability logics can be characterised in terms of semantic properties that are more simple when dealing with Verbrugge frames than in the standard relational setting. Furthermore, some modal logics for interpretability are known to be incomplete w.r.t. Veltman semantics, while being complete w.r.t. Verbrugge semantics.

At the same time, the results collected by Joosten et al. [17] after sever years of active research show that the interactions between Veltman and Verbrugge semantics are subtle: When it comes to models, the two semantics are equally powerful, but when reasoning on frames, the situation is quite different, and variations on the definition of Verbrugge frames lead to different results.

The *main contribution* of the work presented here consists of designing *modular sequent systems* satisfying the main structural desiderata, namely, admissibility

---

4 `supexp` is the function $\lambda x.2^x_x$ with $2^n_0 := n$ and $2^n_{m+1} := 2^{(2^n_m)}$.

5 As explained by Joosten and Visser [18], there is no simple formal definition of "reasonable" theory; in particular, reasonable theories do not have to use the language of arithmetic, so that in the following expression we slightly overload the notation for the sake of conciseness, following similar examples in the literature on this topic.

6 See §3.1 below for a definition of the latter.

of contraction and weakening, invertibility of logical rules, and a *cut-elimination algorithm*.

More precisely, we build on Verbrugge semantics to define labelled sequent calculi for IL and the most relevant of its extensions, including ILM, ILP, and, more interestingly, ILW, which was *not covered before* in proof-theoretic literature. These calculi utilise world and neighbourhood labels to incorporate relevant semantic features into the syntax. At the same time, each proof system adheres to *standard proof-theoretic principles*, ensuring that each connective is addressed by dual left and right rules, supporting a *clear syntactic explanation of its meaning*. With the intent to stick to principled design philosophy, we also defined a non-normal operator $\langle]_x$ for neighbourhoods indexed on world labels, introduced for translating naturally the meaning of the $\rhd$ modality into sequent rules. Moreover, we adapted the methodology used by Negri [28] for *dealing with Noetherianity* of GL-frames, with the intent of capturing within our proof systems the same notion, modulo the necessary generalisation to consider Verbrugge frames for interpretability logics. In short, as Verbrugge semantics is not *per se* expressible as a geometric theory—i.e., cannot be reduced to a finite set of axiom schemas in the language of first-order logic of shape $A \to B$, where A and B do not contain $\forall$ and $\to$—defining sequent rules based on the *geometric rule schemas* by Negri and von Plato [33] was a not-so-easy task, and required a careful approach to the design of the calculi, in order to obtain precise structural properties of the systems for these logics.

By adopting such a principled set of design choices, our calculi markedly exhibit a *modular nature*, where logical rules remain consistent across all systems. In contrast, relational rules for neighbourhood and world labels are introduced to define specific calculi for each considered extension.

In addition to their simplicity and modularity, the calculi boast *robust structural properties*. These include the invertibility of all rules and the admissibility of contraction and cut, proven in a once-for-all way for the base system and its extensions. From the technical viewpoint, we remark that the definition of the cut-elimination algorithm required the use of a *ternary measure of the cut complexity*, which is borrowed from the proof of cut admissibility for G3GL in [28], modulo some tweaks to deal with the more general setting of Verbrugge semantics.

As of the current literature, the comprehensive *proof-theoretic analyses* of interpretability logics presented here are *unparalleled in scope*. The only relevant limitation of the current version of our sequent calculi is that it can be interpreted as a framework "for Veltman frames in disguise" because of the correspondence established by Verbrugge [50] between Veltman models and the version of Verbrugge models we adopted.[7] Still, our findings collected here suggest a promising avenue for the canonical extension of the results presented in the following pages to cover other variants of Verbrugge semantics and more intricate logical frameworks dealing with modal logic for meta-mathematics.

1.3. Paper structure. The paper is organised as follows: In §2, we recall the axiomatic calculi for the main interpretability logics under investigations; in §3, the basic definitions and results in the model theory for interpretability logics are recalled, and the Verbrugge semantics (a.k.a. Generalised Veltman Semantics) (GVS)

[7] We are grateful to an anonymous reviewer for noticing this aspect of our labelled systems.

is discussed according to the most recent literature on the topic. Our family of sequent calculi $\mathsf{G3IL}^\star$ is then presented in §4; §5 is committed to the structural analysis of those calculi, and includes a constructive proof of admissibility of the cut rule for all the extensions. Finally, in §6, we prove the soundness and completeness of our systems w.r.t. the semantic and axiomatic presentations, respectively. A section discussing related and future work closes the article.

§2. Axiomatic calculi. Let us start by recalling the formal definitions of a propositional modal language for interpretability logics.

DEFINITION 2.1. *The set of formulas in a standard formal language $\mathcal{L}_{\Box,\rhd}$ for interpretability logics is given by the following grammar*:

$$p \mid A \wedge B \mid A \vee B \mid A \to B \mid \bot \mid \Box A \mid A \rhd B,$$

where p *belongs to a denumerable set* `Atm` *of atomic propositions. Truth, negation and coimplication are defined in the standard way*: $\top := \bot \to \bot$, $\neg A := A \to \bot$, *and* $A \leftrightarrow B := (A \to B) \wedge (B \to A)$, *respectively. When reasoning in a classical setting, we can also consider a further modal operator* $\Diamond$, *which we define as* $\Diamond A := \neg\Box\neg A$.

The basic axiomatic calculus for interpretability logics is given by the following definition.

DEFINITION 2.2. *Let* $\mathbb{IL}$ *denote the axiomatic system determined by*

- *the axiom schemas of any classical propositional calculus* $\mathbb{CPC}$;
- *schema* K: $\Box(A \to B) \to (\Box A \to \Box B)$;
- *schema* GL: $\Box(\Box A \to A) \to \Box A$;
- *interpretability schemas*[8]
  - IL1: $\Box(A \to B) \to A \rhd B$;
  - IL2: $A \rhd B \to (B \rhd C \to A \rhd C)$
  - IL3: $A \rhd C \to (B \rhd C \to A \vee B \rhd C)$
  - IL4: $A \rhd B \to (\Diamond A \to \Diamond B)$
  - IL5: $\Diamond A \rhd A$
- MP (Modus Ponens) *rule* $\dfrac{A \to B \qquad A}{B}$
- *Necessitation rule* $\dfrac{A}{\Box A}$

As usual, we write $\Gamma \vdash_{\mathbb{IL}} A$ *when* A *is derivable in* $\mathbb{IL}$ *assuming the set of hypotheses* Γ, *with the proviso that the rule of necessitation is applied only to theorems, otherwise the rule is not sound and the deduction theorem fails, as discussed by* [14]. *We write* $\mathbb{IL} \vdash A$ *when* $\Gamma = \varnothing$.

For this calculus, some interesting lemmas are provable.

LEMMA 2.3. *The following hold*

(i) $\mathbb{IL} \vdash \Box\neg A \to (A \rhd B)$;
(ii) $\mathbb{IL} \vdash A \vee \Diamond A \rhd A$;
(iii) $\mathbb{IL} \vdash A \rhd A \wedge \Box\neg A$;

[8] As stated in the §1, here and in the rest of the paper, we assume that $\rhd$ binds stronger than $\to$, but weaker than the other connectives.

(iv) $A \rhd B$, $A \wedge \Box\neg A \rhd B$ *and* $A \rhd B \wedge \Box\neg B$ *are inter derivable over* $\mathbb{IL}$;
(v) $A \rhd \bot$ *and* $\Box\neg A$ *are inter derivable over* $\mathbb{IL}$;
(vi) $\mathbb{IL} \vdash \Diamond A \rhd \neg(A \rhd \Diamond A)$.

Proof. Refer to e.g., [54]. □

Item (v) of the previous lemma shows that we could dismiss the □ modality since we could define $\Box A$ as $\neg A \rhd \bot$. This definition would simplify our base language. However, we will see in §3.1, §3.2, and, more poignantly, §4.1, that it is possible to define □ via ▷ as well as ▷ via □ by semantics considerations. Those will lead to the design of our family of sequent calculi in §4.[9]

2.1. Axiomatic extensions. On top of $\mathbb{IL}$, it is possible to add further modal principles that have specific relevance for arithmetical realisations. Here, we will consider the following calculi.

Definition 2.4. *Let us define as proper extensions of* $\mathbb{IL}$

- $\mathbb{ILM} := \mathbb{IL} + \mathsf{M}$, *where*
$$\mathsf{M} := A \rhd B \to A \wedge \Box C \rhd B \wedge \Box C$$
is called the Montagna schema;
- $\mathbb{ILP} := \mathbb{IL} + \mathsf{P}$, *where*
$$\mathsf{P} := A \rhd B \to \Box(A \rhd B)$$
is called the persistence schema;
- $\mathbb{ILW} := \mathbb{IL} + \mathsf{W}$, *where*
$$\mathsf{W} := A \rhd B \to A \rhd B \wedge \Box\neg A;$$
- $\mathbb{ILKM}1 := \mathbb{IL} + \mathsf{KM1}$, *where*
$$\mathsf{KM1} := A \rhd \Diamond B \to \Box(A \to \Diamond B);$$
- $\mathbb{ILM}_0 := \mathbb{IL} + \mathsf{M}_0$, *where*
$$\mathsf{M}_0 := A \rhd B \to \Diamond A \wedge \Box C \rhd B \wedge \Box C.$$

§3. Semantics for interpretability logics. As stated in the introduction, we know—after [2, 40], and [54, 57], respectively—that $\mathbb{ILM}$ is the interpretability logic for arithmetical theories proving full induction, and $\mathbb{ILP}$ captures the properties of formal interpretability over finitely axiomatised theories proving the totality of the superexponential function.[10]

Their proofs are based on Solovay's strategy for arithmetical completeness of Gödel–Löb logic as exposed in [5, 43]; therefore, they effectively use the characterisation of

[9] Similar considerations underlie the definition of the tableaux systems in [15] .
[10] Since $\mathbb{ILM}$ and $\mathbb{ILP}$ are not contained in and do not contain each other, those completeness results entail, for instance, that Montagna's principle holds for Peano arithmetic but does not hold for Gödel–Bernays set theory; on the contrary, the persistence principle holds for Gödel–Bernays set theory, but does not hold for Peano arithmetic. Nevertheless, they share the same provability logic, namely, GL.

those axiomatic systems in terms of relational models. We now turn our presentation to these semantic aspects of interpretability logics.

3.1. Veltman semantics. To obtain a standard relational semantics for interpretability logics from possible world semantics one usually proceeds by "decorating" frames with additional indexed accessibility relations.

Definition 3.1. *A Veltman frame $\mathcal{F}$ consists of*:

- *a nonempty set W of possible worlds*;
- *a binary relation R on W which is transitive and Noetherian*[11] ;
- *a collection $\{S_x \mid x \in W\}$ of binary relations which are reflexive, transitive and such that*
  - *if yS_xz, then $y \in R[x]$ and $z \in R[x]$, where $R[x] := \{y \in W \mid xRy\}$; and*
  - *if $xRyRz$, then yS_xz.*

A Veltman model $\mathcal{M}$ *is obtained by adding an evaluation function $v : \texttt{Atm} \times W \to \{0,1\}$ to a given Veltman frame, as usual. A* forcing relation $\Vdash$ *is then obtained by a standard definition for propositional connectives and $\Box$-modality*

$$x \Vdash_{\mathcal{M}} \Box B \quad \text{iff} \quad \text{for all } y \in W, \text{if } xRy \text{ then } y \Vdash_{\mathcal{M}} B$$

while for $\rhd$-modality we stipulate that

$$x \Vdash_{\mathcal{M}} A \rhd B \quad \text{iff} \quad \text{for all } y \in W, \text{ if } xRy \text{ and } y \Vdash_{\mathcal{M}} A, \text{then there exists a } z \in W \text{ such that } yS_xz \text{ and } z \Vdash_{\mathcal{M}} B.$$

We write $\vDash_{\mathcal{M}} A$ when A is forced by any world in $\mathcal{M}$; similarly, we write $\vDash_{\mathcal{F}} A$ when $\vDash_{\mathcal{M}} A$ for any model $\mathcal{M}$ based on $\mathcal{F}$.

For extensions, some frame conditions are needed. A *frame condition* for a modal schema A is a (first or higher order) formula (A) in the language $\{W, R, \{S_x\}\}$ such that the structure $\mathcal{F}$ satisfies the property (A) if and only if $\vDash_{\mathcal{F}} A$.

In [53, 54], many principles for interpretability were proposed first, together with their semantic characterisations. We then know that

$$(\mathsf{W}) = (\mathsf{KW1}) = (\mathsf{F}) = \text{"}R \circ S_x \text{ is Noetherian" for all } x \in W,$$

where

$$\begin{array}{lcl} \mathsf{KW1} & := & A \rhd \Diamond\top \to \top \rhd \neg A \\ \mathsf{F} & := & A \rhd \Diamond A \to \Box\neg A. \end{array}$$

Moreover we have that

$$(\mathsf{M}) = (\mathsf{KM1}) = (\mathsf{KM2}) = \text{ if } yS_xzRu, \text{ then } yRu,$$

where

$$\mathsf{KM2} \quad := \quad A \rhd B \to (\Box(B \to \Diamond C) \to \Box(A \to \Diamond C)),$$

and we know, after Visser [53], that KM1 and KM2 are interderivable over $\mathbb{IL}$.

By using Veltman semantics, Švejdar [46] proved that

$$\mathbb{IL}\{\mathsf{F}, \mathsf{KW1}\} \nvdash \mathsf{KW1}^\circ,$$

[11] We say that a relation R is Noetherian on W when for any $X \subseteq W$ there exists a $w \in X$ such that, for no $x \in X$, wRx.

where

$$\mathsf{KW1}^\circ := A \wedge B \rhd \Diamond A \to A \rhd (A \wedge \neg B).$$

After the work on those systems by Švejdar [46], we furthermore know that $\mathbb{ILF}$,$\mathbb{ILKW}1$, $\mathbb{ILKW}1^\circ$ are *incomplete* w.r.t. the standard relational semantics.

3.2. Verbrugge semantics. Verbrugge semantics (also called GVS) rescues the situation. Rineke Verbrugge has developed it in [50] by considering interpretations reminiscent of neighbourhood semantics for non-classical logics.

To be more precise, each S_x is now a relation between worlds and *sets of worlds*, satisfying specific properties identified by the schemas for $\rhd$.

DEFINITION 3.2. *A Verbrugge frame $\mathcal{F}$ consists of*

- *a finite set $W \neq \varnothing$;*
- *a binary relation $R \subseteq W \times W$ which is irreflexive and transitive;*
- *a W-indexed set of relations $S_x \subseteq R[x] \times (\wp(R[x]) \setminus \{\varnothing\})$;*

satisfying the following conditions:

- *Quasi-reflexivity*: *if xRy then $yS_x\{y\}$;*
- *Definiteness*: *if $xRyRz$ then $yS_x\{z\}$;*
- *Monotonicity*: *if yS_xa and $a \subseteq b \subseteq R[x]$ then yS_xb;*
- *Quasi-transitivity*: *if yS_xa and vS_xb_v for all $v \in a$, then $yS_x(\bigcup_{v\in a} b_v)$.*

REMARK 1. *The frames used in Verbrugge semantics are, in a sense, relational frames for Gödel–Löb logic enriched with indexed relations between worlds and nonempty sets of worlds.*[12] *Those frames for Gödel–Löb logic are conventionally those that are endowed with an accessibility relation that is transitive and Noetherian. However, we know that a semantics based on irreflexive, transitive* and *finite frames (ITF) is also adequate for Gödel–Löb logic. Therefore, we adopt here this second semantic characterization of the system for provability, noting that this has no impact on the "neighbourhood aspects" of Verbrugge semantics for the logics we investigate here, since they all satisfy the finite model property for GVS—and the same could have been done for standard Veltman semantics.*[13]

A Verbrugge model *is obtained by considering a usual evaluation function, which can be extended to a* forcing relation *defined as for standard semantics, with only one difference*:

$$x \Vdash A \rhd B \text{ iff for all } y \text{ if } xRy \text{ and } y \Vdash A, \text{ then there exists an } a \text{ such that } yS_xa \text{ and } a \Vdash^\forall B,$$

where $a \Vdash^\forall B$ abbreviates the expression "for any $z \in a, z \Vdash B$".

As for relational semantics, extensions for $\mathbb{IL}$ need *generalised frame conditions*: We denote by $(A)_{gen}$ the frame condition w.r.t. Verbrugge semantics corresponding to the modal schema A.

---

12 We choose to refer to the latter as neighbourhoods, following the standard terminology in neighbourhood semantics for non-classical logics.

13 As the reviewers pointed out, our current methodology could not be implemented on extensions of IL that do not satisfy the finite model property.

We know that the following hold[14]

$$
\begin{array}{lll}
(\mathsf{M})_{gen} & = & yS_xa \Rightarrow \exists b \subseteq a, yS_xb \,\&\, R[b] \subseteq R[y] \\
(\mathsf{KM1})_{gen} & = & yS_xa \Rightarrow \exists i \in a, \forall z(iRz \Rightarrow yRz) \\
(\mathsf{P})_{gen} & = & xRx'RyS_xa \Rightarrow \exists b \subseteq a, yS_{x'}b \\
(\mathsf{M}_0)_{gen} & = & wRuRxS_wa \Rightarrow \exists b \subseteq a, uS_wb \,\&\, R[b] \subseteq R[u] \\
(\mathsf{P}_0)_{gen} & = & wRxRuS_wa \,\&\forall v \in a(R[v] \cap b \neq \varnothing) \Rightarrow \exists c \subseteq b, uS_xc \\
(\mathsf{W})_{gen} & = & yS_xa \Rightarrow \exists b \subseteq a, yS_xb \,\&\, R[b] \cap S_x^{-1}b = \varnothing \\
(\mathsf{R})_{gen} & = & wRxRuS_wa \Rightarrow \forall c \in \mathcal{C}(x,u), \exists b \subseteq a, xS_Wb \,\&\, R[b] \subseteq c \\
(\mathsf{W}^*)_{gen} & = & (\mathsf{M}_0)_{gen} \,\&\, (\mathsf{W})_{gen},
\end{array}
$$

where

- $i \in R[b]$ iff there is an $x \in b$ such that xRi;
- $i \in S_x^{-1}b$ iff iS_xb;
- $\mathcal{C}(x,u) := \{c \subseteq R[x] \mid \forall d, uS_xd \Rightarrow d \cap c \neq \varnothing\}$;
- $\mathsf{P}_0 := A \rhd \Diamond B \to \Box(A \rhd B)$;
- $\mathsf{R} := A \rhd B \to \neg(A \rhd \neg C) \rhd B \wedge \Box C)$;
- $\mathsf{W}^* := A \rhd B \to B \wedge \Box C \rhd B \wedge \Box C \wedge \Box\neg A$.

For the basic system $\mathbb{IL}$, completeness results are known w.r.t. both standard Veltman semantics and GVS. Moreover, the techniques used to prove the completeness theorem for that system w.r.t. GVS can be easily extended to consider analogous results for $\mathbb{ILM}$, $\mathbb{ILP}$, and $\mathbb{ILW}$ proven by de Jongh and Veltman [7, 8] w.r.t. ordinary Veltman semantics.

However, for the other extensions, the proof of modal completeness can be quite convoluted and very sensitive to the logic under consideration, so proving that an extension of a given system is complete may need a very different proof strategy w.r.t. the one used for the completeness of the original subsystem. Some promising advances have been made recently by Joost Joosten and collaborators in a series of works aiming at developing a modular and uniform methodology to deal with GVS for interpretability logics: The most recent literature on the topic includes Mikec and Vukovic [27], who leverage techniques from Bilkova et al. [3] and Goris et al. [11], subsequently surveyed by Joosten et al. [17] and Rovira et al. [38].[15]

In any case, investigations on GVS suffice to establish that, for the interpretability logics, we have mentioned the interdependencies rendered in Figure 1 hold.

Finally, we can summarise the current model-theoretic knowledge on interpretability logic by the glossary in Figure 2

§4. Design of the labelled sequent calculi. We have seen that the language $\mathcal{L}_{\Box,\rhd}$ is somehow redundant: After Lemma 2.3(v), we know that $\Box A$ is equivalent to $\neg A \rhd \bot$. This equivalence invites to minimise the formal language for interpretability by considering the $\rhd$-modality as primitive. The resulting language will be denoted by $\mathcal{L}_{\rhd}$.

We need now to rephrase the inductive definition of well-formed formulas of $\mathbb{IL}$—and its extensions—as follows.

[14] We need to use symbolic connectives for the meta-level to enhance readability.

[15] It is worth noticing that there exists flourishing research in finding even more general interpretability principles, whose semantics is still under investigation (see, e.g., [23, 24].

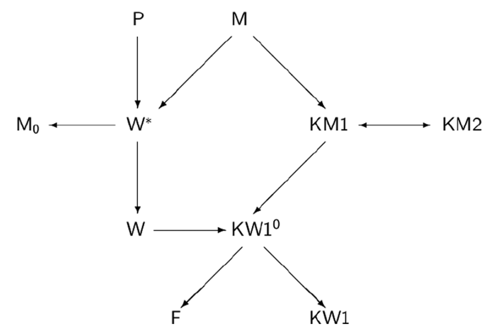

Figure 1. Interdependencies among interpretability logics, after Vuković [56]. An arrow from S to S′ is interpreted as $\mathbb{IL} + \mathsf{S}' \subsetneq \mathbb{IL} + \mathsf{S}$.

Modal principle	Veltman completeness	GVS completeness	Veltman FMP	GVS FMP
M	✓	✓	✓	✓
P	✓	✓	✓	✓
W	✓	✓	✓	✓
W*	✓	✓	?	✓
M_0	✓	✓	?	✓
P_0	×	✓	?	✓
R	?	✓	?	✓
F	×	?	✓	✓

Figure 2. This information is collected from the results in [50], [7, 8], [12], [26, 27], [37]. Here FMP abbreviates "finite model property".

Definition 4.1. *The set of well-formed formulas of* $\mathbb{IL}$ *and its extensions w.r.t.* $\mathcal{L}_\rhd$ *is given by the following grammar*

$$\mathtt{Form}_\rhd ::= \ p \mid A \wedge B \mid A \vee B \mid A \to B \mid \bot \mid A \rhd B \ ,$$

where $p \in \mathtt{Atm}$ *and* $A, B \in \mathtt{Form}_\rhd$.

As the reader might expect, we define $\neg A := A \to \bot$, $A \leftrightarrow B := (A \to B) \wedge (B \to A)$, $\Box A := \neg A \rhd \bot$, and $\Diamond A := \neg\Box\neg A$.

On the axiomatic side, such a minimalist choice about the basic modal language (only $\rhd$ is used as modal operator) is reflected by a minimalist axiomatisation of the basic modal system for interpretability.

Definition 4.2. *Let* $\mathbb{IL}_\rhd$ *denote the axiomatic calculus defined in* [55] *by*[16]

- *Axiom schemas of* $\mathbb{CPC}$;
- *schema* IL2: $A \rhd B \to (B \rhd C \to A \rhd C)$;
- *schema* IL3: $A \rhd C \to (B \rhd C \to A \vee B \rhd C)$;
- *schema* IL-Löb: $A \rhd (A \wedge (A \rhd \bot))$;
- MP *Rule* $\dfrac{A \to B \qquad A}{B}$;
- $\rhd$ *Rule* $\dfrac{A \to B}{A \rhd B}$.

[16] Recall from the footnote 8 that we assume that $\rhd$ binds stronger than $\to$, but weaker than the other connectives.

The extensions of $\mathbb{IL}_{\triangleright}$ are obtained by adding the axiom schemas that we discussed in §2.1 *and* §3; *they will be denoted analogously to the systems based on multimodal $\mathbb{IL}$.*

The calculi we will present are obtained by labelling the formulas in $\texttt{Form}_{\triangleright}$. The classes of models we are using for the definition of these calculi are based on Verbrugge semantics: This is in line with the procedure initiated by Negri [28] for relational semantics and generalised further to neighbourhood semantics in e.g., [9, 10, 31].[17]

4.1. Core system. By G3IL we denote the labelled sequent calculus for $\mathbb{IL}_{\triangleright}$. The design of G3IL is based on an explicit internalisation of GVS employing labels, which formalises the semantic information into a proof system.

According to Definition 3.2, the forcing condition for the $\triangleright$-modality is

$$(\sharp) \quad x \Vdash A \triangleright B \quad \text{iff} \quad \text{for all } y, \text{if } xRy \text{ and } y \Vdash A, \text{ then there exists an } a \text{ such that } yS_x a \text{ and } a \Vdash^{\forall} B.$$

As it comes, that forcing condition cannot be directly translated into a single sequent calculus geometric rule because of the presence of alternating nested quantifiers on the right-hand side of ($\sharp$).[18]

To circumvent that issue, it is convenient to introduce an intermediate indexed modality, which obeys the following forcing condition

$$(\flat) \quad y \Vdash \langle]_x B \quad \text{iff} \quad \text{there exists an } a \text{ such that } yS_x a \text{ and } a \Vdash^{\forall} B,$$

where $a \Vdash^{\forall} B$ abbreviates the expression "for any $z \in a, z \Vdash B$", as in ($\sharp$).

The forcing condition for $\triangleright$ can then be rephrased as

$$(\sharp\flat) \quad x \Vdash A \triangleright B \quad \text{iff} \quad \text{for all } y, \text{if } xRy \text{ and } y \Vdash A, \text{ then } y \Vdash \langle]_x B.$$

Unfortunately, this is not enough yet. Models for $\mathbb{IL}$ are based on frames for Gödel–Löb logic, that is: irreflexive, transitive and *finite*—or, alternatively, transitive and Noetherian. Neither finiteness nor Noetherianity can be expressed by a semantic rule in line with the methodology of explicit internalisation available so far in the literature.

However, the treatment of Gödel–Löb logic described in [30] gives the right hint for proceeding with the design. Notice first that condition ($\sharp\flat$) establishes the logical equivalence

$$x \Vdash A \triangleright B \qquad \text{iff} \qquad x \Vdash \Box(A \to \langle]_x B). \tag{1}$$

Moreover, we know that in any model based on *ITF*[19]

$$x \Vdash \Box A \qquad \text{iff} \qquad \text{for any } y,\ \text{if } xRy \text{ and } y \Vdash \Box A,\ \text{then } y \Vdash A. \tag{2}$$

---

17 Such a move is possible in virtue of the adequacy results for the systems under investigations w.r.t. GVS that we briefly recalled in the glossary at the end of §3.2.

18 Recall from e.g., [52], that a geometric formula is a formula in the language of first-order logic of shape $A \to B$, where A and B do not contain $\forall$ and $\to$. Geometric rules for sequent calculi are discussed in details by Negri and von Plato [32].

19 By *ITF*, it is common to denote the class of relational frames that are irreflexive, transitive, and finite. Refer also to Remark 1 above. Notice, however, that the same equivalence holds in frames that are only Noetherian and transitive.

This suggests to index with worlds the $\rhd$-modality and, by chaining the equivalences (1) and (2), we can safely take the following forcing condition for that indexed $\rhd$-modality whenever we are reasoning in models based on GVS:

$$(\natural) \quad x \Vdash A \rhd_i B \quad \text{iff} \quad \begin{array}{l} \text{for all } y, \text{if } xRy \text{ and } y \Vdash A \rhd_i B, \\ \text{then, if } y \Vdash A,\ y \Vdash \langle]_i B. \end{array}$$

Let $\texttt{Form}^i_{\rhd}$ denote the formulas allowing indexed $\rhd$-modalities as well as the intermediate indexed modalities $\langle]_x$.

We are now ready to define the labelled sequent calculus G3IL.

Definition 4.3. *Let $i, j, k, \dots, x, y, z, \dots$ be variables for worlds in a Verbrugge model, and $s, t, u, \dots$ variables for neighbourhoods. By* neighbourhood term *$a, b, c, \dots$ we refer to a neighbourhood variable or an expression of shape $\{x\}$ or $R[x]$ for* x *a world variable.*

Relational atoms *are formulas of the following form and meaning*:

- *$y \in R[x]$, "world y is accessible to world x"*;
- *$yS_x a$, "neighbourhood a is S_x-accessible to world y"*;
- *$y \in a$, "world y is a member of neighbourhood a"*;
- *$a \subseteq b$, "neighbourhood a is included into neighbourhood b".*

We will use $\{x\}$ to denote the singleton neighbourhood of exactly the world x.

Labelled formulas *are defined as follows, for $A \in \texttt{Form}^i_{\rhd}$*:

- *Relational atoms are labelled formulas*;
- *$x : A$, "world x forces formula A"*;
- *$a \Vdash^{\forall} A$, "formula A is forced by any world belonging to neighbourhood a".*

Definition 4.4. *Sequents of* G3IL *are expressions $\Gamma \Rightarrow \Delta$, where Γ and Δ are multisets of labelled formulas, and relational atoms may occur* only *in Γ.*

The rules defining G3IL are given in Figure 3.

Remark 2. *Some of those rules might deserve a little explanation*:

- *Side condition $(x!)$ in $\mathcal{R} \Vdash^{\forall}$ expresses the fact that* x *is a 'fresh varibale', i.e., it does not occur in the conclusion of the rule; similarly for $(y!)$ in $\mathcal{R}\rhd_i$; the meaning of $(a!)$ in $\mathcal{L}\langle]$ is analogous; the meaning of $(z!)$ in the rule* NE *is analogous.*
- *The rules for $\langle]_x$ are defined according to the forcing condition $(\flat)$, following the standard practice for labelled sequent calculi based on neighbourhood semantics.*[20]
- *The rules for $\rhd_i$ are defined according to the forcing condition $(\natural)$.*
- *The rules for GVS are defined as geometric rules of Negri and von Plato [33]; in particular, we opted for an alternative definition of quasi-transitivity of the indexed S-relation. The condition imposed by Definition 3.2 is*

$$\textit{if } yS_x a \textit{ and } vS_x b_v \textit{ for all } v \in a, \textit{ then } yS_x\left(\textstyle\bigcup_{v \in a} b_v\right).$$

This condition cannot be directly translated as a geometric rule with the language at hand: recall from e.g., [33] that every geometric formula can be shown to be equivalent to a formula of shape

$$\forall \vec{x}(P_1 \wedge \dots \wedge P_m \to \exists \vec{y}_1 \mathbf{Q_1} \vee \dots \vee \exists \vec{y}_n \mathbf{Q_n}),$$

[20] Refer to e.g., [9, 31].

Initial sequents

$x : p, \Gamma \Rightarrow \Delta, x : p$

$x : A \rhd_i B, \Gamma \Rightarrow \Delta, x : A \rhd_i B$

Classical propositional rules: the usual ones for e.g. G3K in Negri (2005).

Local forcing rules

$$\frac{x : A, x \in a, a \Vdash^{\forall} A, \Gamma \Rightarrow \Delta}{x \in a, a \Vdash^{\forall} A, \Gamma \Rightarrow \Delta}\ \mathcal{L}\Vdash^{\forall}$$

$$\frac{x \in a, \Gamma \Rightarrow \Delta, x : A}{\Gamma \Rightarrow \Delta, a \Vdash^{\forall} A}\ \mathcal{R}\Vdash^{\forall}_{(x!)}$$

Intermediate modality rules

$$\frac{yS_x a, a \Vdash^{\forall} A, \Gamma \Rightarrow \Delta}{y : \langle]_x A, \Gamma \Rightarrow \Delta}\ \mathcal{L}\langle]_{(a!)}$$

$$\frac{yS_x a, \Gamma \Rightarrow \Delta, y : \langle]_x A, a \Vdash^{\forall} A}{yS_x a, \Gamma \Rightarrow \Delta, y : \langle]_x A}\ \mathcal{R}\langle]$$

Interpretability modality rules

$$\frac{y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta, y : A \qquad y : \langle]_i B, y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta \qquad y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta, y : A \rhd_i B}{y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta}\ \mathcal{L}\rhd_i$$

$$\frac{y \in R[x], y : A, y : A \rhd_i B, \Gamma \Rightarrow \Delta, y : \langle]_i B}{\Gamma \Rightarrow \Delta, x : A \rhd_i B}\ \mathcal{R}\rhd_{i(y!)}$$

Rules for GVS

$$\frac{a \subseteq a, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}\ \mathit{Refl}\subseteq \qquad \frac{a \subseteq c, a \subseteq b, b \subseteq c, \Gamma \Rightarrow \Delta}{a \subseteq b, b \subseteq c, \Gamma \Rightarrow \Delta}\ \mathit{Trans}\subseteq$$

$$\frac{x \in b, x \in a, a \subseteq b, \Gamma \Rightarrow \Delta}{x \in a, a \subseteq b, \Gamma \Rightarrow \Delta}\ \mathcal{L}\subseteq$$

$$\frac{x \in \{x\}, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}\ \mathit{Sing}$$

$$\frac{Atm(y), Atm(x), y \in \{x\}, \Gamma \Rightarrow \Delta}{Atm(x), y \in \{x\}, \Gamma \Rightarrow \Delta}\ \mathit{Repl}_1 \qquad \frac{Atm(x), Atm(y), y \in \{x\}, \Gamma \Rightarrow \Delta}{Atm(y), y \in \{x\}, \Gamma \Rightarrow \Delta}\ \mathit{Repl}_2$$

where $Atm(x)$ has one of the following forms: $x : p, x \in a, x \in \{z\}, x \in R[z], z \in R[x], xS_z a, zS_x a$.

$$\frac{}{x \in R[x], \Gamma \Rightarrow \Delta}\ \mathit{Irrefl} \qquad \frac{z \in R[x], y \in R[x], z \in R[y], \Gamma \Rightarrow \Delta}{y \in R[x], z \in R[y], \Gamma \Rightarrow \Delta}\ \mathit{Trans}$$

$$\frac{z \in a, yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ \mathit{NE}_{(z!)} \qquad \frac{y \in R[x], a \subseteq R[x], yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ \mathit{DefS1}$$

$$\frac{yS_x\{z\}, y \in R[x], z \in R[y], \Gamma \Rightarrow \Delta}{y \in R[x], z \in R[y], \Gamma \Rightarrow \Delta}\ \mathit{DefS2} \qquad \frac{yS_x b, yS_x a, a \subseteq b, b \subseteq R[x], \Gamma \Rightarrow \Delta}{yS_x a, a \subseteq b, b \subseteq R[x], \Gamma \Rightarrow \Delta}\ \mathit{Mono}$$

$$\frac{yS_x\{y\}, y \in R[x], \Gamma \Rightarrow \Delta}{y \in R[x], \Gamma \Rightarrow \Delta}\ \mathit{Qrefl} \qquad \frac{yS_x b, yS_x a, z \in a, zS_x b, \Gamma \Rightarrow \Delta}{yS_x a, z \in a, zS_x b, \Gamma \Rightarrow \Delta}\ \mathit{Qtrans6}$$

Figure 3. Rules for G3IL.

where P_i and $\mathbf{Q_j}$ are atomic formulas and finite conjunctions of atomic formulas, respectively.Every geometric formula can then be converted into an n*-ary geometric rule*:

$$\frac{Q_{11}[y_1/x_1], \dots, Q_{1k_1}[y_1/x_1], \vec{P}, \Gamma \Rightarrow \Delta \qquad \cdots \qquad Q_{n1}[y_n/x_n], \dots, Q_{nk_m}[y_n/x_n], \vec{P}, \Gamma \Rightarrow \Delta}{\vec{P}, \Gamma \Rightarrow \Delta}\ Geom,$$

where $\vec{P}$ corresponds to $P_1, \dots, P_m$ and $[y_i/x_i]$ denotes the replacement of the variable x_i with a fresh variable y_i which does not occur in the conclusion of the rule.Nevertheless, in the literature, it is possible to find several different conditions for quasi-transitivity:[21]

Nr.	Semantic requirement for transitivity
(1)	$uS_xY \Rightarrow \forall \{Y_y\}_{y\in Y}\Big((\forall\, y \in Y\ yS_xY_y) \Rightarrow \exists Z \subseteq \bigcup_{y\in Y} Y_y\ uS_xZ\Big)$
(2)	$uS_xY \Rightarrow \forall \{Y_y\}_{y\in Y}\Big((\forall\, y \in Y\ yS_xY_y) \Rightarrow uS_x \bigcup_{y\in Y} Y_y\Big)$
(3)	$uS_xY \Rightarrow \exists\, y \in Y\, \forall Y'(yS_xY' \Rightarrow \exists\, Y'' \subseteq Y'\ uS_xY'')$
(4)	$uS_xY \Rightarrow \exists\, y \in Y\, \forall Y'(yS_xY' \Rightarrow uS_xY')$
(5)	$uS_xY \Rightarrow \forall\, y \in Y\, \forall Y'(yS_xY' \Rightarrow \exists\, Y'' \subseteq Y'\ uS_xY'')$
(6)	$uS_xY \Rightarrow \forall\, y \in Y\, \forall Y'(yS_xY' \Rightarrow uS_xY')$
(7)	$uS_xY \Rightarrow \forall\, y \in Y\, \forall Y'(yS_xY'\ \&\ y \notin Y' \Rightarrow \exists\, Y'' \subseteq Y'\ uS_xY'')$
(8)	$uS_xY \Rightarrow \forall\, y \in Y\, \forall Y'(yS_xY'\ \&\ y \notin Y' \Rightarrow uS_xY')$

Condition 2 *is the most natural*: *the monotone closure of any S_x that satisfies any of conditions* 1–8 *also satisfies condition* 2, *allowing us to define an equivalent model. For conditions* 3–6, *obtaining a related standard model from a GVS model is always possible, as demonstrated in* [17, 50]. *The rule* Qtrans6 *represents the natural formalization of condition* 6 *from the previous table, which is the simplest one according to the methodology outlined in* [33] *that we have briefly summarised. However, the correspondence between GVS models satisfying condition* 6 *and standard Veltman models limits the expressive power of our current systems. To achieve full adequacy of the interpretability logics under investigation with respect to the notion of Verbrugge frames standard in the literature—i.e., based on condition* 2*—we need to revise this semantic rule, as issue that we intend to tackle in future work*.[22]

- *Rule* Sing *assures that the singleton contains at least one element*; *rules* Repl_1 *and* Repl_2 *that it contains at most one element, for indiscernibility of identicals.*
- *From Figure* 3 *the rules obtained by the closure condition of the system, discussed in* [28], *are omitted. For some rules dealing with GVS—for instance,* Trans*— there might be a duplication of a relational atom in the conclusion. Structural considerations—namely, the desideratum of admissibility of contraction—require then that a new rule is added to the system, in which the duplicated formulas are contracted into one.*[23] *However, the rules added to a system in order to satisfy such*

21 The table is taken from [17].

22 See our Remark 5 and §7 below for further discussion of this source of limitations.

23 It is relevant to notice here that for each semantic characterisation, there is only a bounded number of additional rules generated by the closure condition. Moreover, that number is generally made smaller since many cases of contracted rules are shown to be admissible in the base calculus. Refer to [33, chaps. 6 and 11] for an exhaustive description of the procedure and its relevance for labelled calculi for modal logics. Similar considerations hold for the rules in Figure 4.

Additional rules for GVS

$$\frac{x \in a, y \in R[x], y \in R[a], \Gamma \Rightarrow \Delta}{y \in R[a], \Gamma \Rightarrow \Delta}\ Rset1_{(x!)} \qquad \frac{y \in R[a], x \in a, y \in R[x], \Gamma \Rightarrow \Delta}{x \in a, y \in R[x], \Gamma \Rightarrow \Delta}\ Rset2$$

$$\frac{yS_x a, y \in S_x^{-1}a, \Gamma \Rightarrow \Delta}{y \in S_x^{-1}a, \Gamma \Rightarrow \Delta}\ Sset1 \qquad \frac{y \in S_x^{-1}a, yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ Sset2$$

$$\frac{c \subseteq a, c \subseteq b, c \subseteq a \cap b, \Gamma \Rightarrow \Delta}{c \subseteq a \cap b, \Gamma \Rightarrow \Delta}\ \cap_1 \qquad \frac{c \subseteq a \cap b, c \subseteq a, c \subseteq b, \Gamma \Rightarrow \Delta}{c \subseteq a, c \subseteq b, \Gamma \Rightarrow \Delta}\ \cap_2$$

$$\frac{}{x \in \varnothing, \Gamma \Rightarrow \Delta}\ \mathcal{L}\varnothing$$

Rules for interpretability principles

$$\frac{b \subseteq a, yS_x b, R[b] \subseteq R[y], yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ M_{(b!)} \qquad \frac{z \in a, R_z \subseteq R[y], yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ KM1_{(z!)}$$

$$\frac{b \subseteq a, zS_y b, y \in R[x], z \in R[y], zS_x a, \Gamma \Rightarrow \Delta}{y \in R[x], z \in R[y], zS_x a, \Gamma \Rightarrow \Delta}\ P_{(b!)} \qquad \frac{b \subseteq a, yS_x b, R[b] \cap S_x^{-1}a \subseteq \varnothing, yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ W_{(b!)}$$

$$\frac{b \subseteq a, yS_x b, R[b] \subseteq R[y], y \in R[x], z \in R[y], zS_x a, \Gamma \Rightarrow \Delta}{y \in R[x], z \in R[y], zS_x a, \Gamma \Rightarrow \Delta}\ M_{0_{(b!)}}$$

Figure 4. Rules for G3IL*.

a closure condition play no role in the proof of semantic completeness of the calculi we are considering here, and they can be shown to be admissible. This justifies our omission in favour of a better readability of the figure.

Here, we see that the version of Verbrugge semantics for $\mathbb{IL}$ *adopted here can be considered a geometric theory, and thus, it can be formalised by a sequent calculus based on purely geometric rules.*

4.2. Extensions. Calculi for extensions of $\mathbb{IL}$ are denoted by adding to G3IL the name of a modal schema as apex. Thus, for instance, G3IL$^{\mathsf{M}}$ is the labelled calculus for $\mathbb{ILM}$, and G3IL$^{\mathsf{P}}$ is the labelled calculus for $\mathbb{ILP}$. We denote by G3IL* the whole family of calculi for the interpretability logics considered in Definition 2.4.

Figure 4 shows the rules for the $\mathbb{IL}$ extensions in which we are interested.

As the reader sees, these rules are obtained by considering the generalised frame conditions characterising each $\mathbb{IL}$ extension. Moreover, we need to consider an extension of the language of labelled formulas.

Definition 4.5. *Extend Definition 4.3 by considering among* neighbourhood labels *expressions of shape* $\varnothing$, $a \cap b$, $R[a]$, *and* $S_x^{-1}a$ *with* a, b *neighbourhood labels and* x *a world label*. Relational atoms *are accordingly defined as follows*:

- $y \in a \cap b$, *"world* y *if a member of both neighbourhood* a and *neighbourhood* b*"*;
- $y \in R[a]$, *"world* y *is accessible to a world* x *belonging to* a*"*;
- $y \in S_x^{-1}a$, *"neighbourhood* a *is* S_x*-accessible to a world* y*"*;
- $x \in \varnothing$, *"world* x *is a member of the neighbourhood* $\varnothing$*" – we take the latter as the only* constant for neighbourhoods *of our language.*

Labelled formulas *are defined as in Definition 4.3 w.r.t. this extended set of relational atoms.*

§5. Structural properties. We now want to study the structural properties of G3IL*. In order to proceed, we need first some preliminary definitions.

By the *height* of a derivation, we mean the number of nodes occurring in the longest derivation branch minus one. In particular, the height of a derivation consisting only of an initial sequent is 0. We write $\overset{n}{\vdash} \Gamma \Rightarrow \Delta$ whenever there is a derivation of the sequent $\Gamma \Rightarrow \Delta$ in G3IL* with height bounded by n.

Next, we need the notion of *weight* of labelled formulas.

DEFINITION 5.1. *The weight of relational atoms is* 0. *As for the other labelled formulas, let us say that the label of* $x : A$ *is* x, *and the label of* $a \Vdash^{\forall} A$ *is* a. *The label of a formula* φ *is denoted by* $l(\varphi)$, *and* $p(\varphi)$ *denotes the pure part of the formula, i.e., the part of the formula without the label and the forcing condition.*

The weight $\mathfrak{w}(\varphi)$ *of a labelled formula* φ *which is not a relational atom is given by the ordered pair* $\langle \mathfrak{w}(p(\varphi)), \mathfrak{w}(l(\varphi)\rangle$, *where*

- *For all world labels* x *and all neighbourhood labels* a, $\mathfrak{w}(x) = 0$ *and* $\mathfrak{w}(a) = 1 + n(\cap)$, *where* $n(\cap)$ *denotes the number of formal intersections in* a;
- $\mathfrak{w}(p) = \mathfrak{w}(\bot) = 1$;
- $\mathfrak{w}(A \circ B) = \mathfrak{w}(A) + \mathfrak{w}(B) + 1$, *for* $\circ$ *conjunction, disjunction or implication*;
- $\mathfrak{w}(\langle]_i A) = \mathfrak{w}(A) + 1$;
- $\mathfrak{w}(A \rhd_i B) = \mathfrak{w}(A) + \mathfrak{w}(B) + 2$.

In the following, when reasoning by induction on the weight of a labelled formula we do so by considering the lexicographic order on $\langle \mathfrak{w}(p(\varphi)), \mathfrak{w}(l(\varphi)\rangle$:
$\langle \mathfrak{w}(p(\varphi)), \mathfrak{w}(l(\varphi)\rangle \leq \langle \mathfrak{w}(p(\varphi')), \mathfrak{w}(l(\varphi')\rangle$ *if and only if*
$\mathfrak{w}(p(\varphi)) < \mathfrak{w}(p(\varphi'))$ *or*
$\mathfrak{w}(p(\varphi)) = \mathfrak{w}(p(\varphi'))$ *and* $\mathfrak{w}(l(\varphi) < \mathfrak{w}(l(\varphi')$

For substitution of labels, we can rely on the definitions given in [31, 33]. We borrow notation from those works and write, e.g., $(a \Vdash^{\forall} A)[b/a]$ to mean the result of simultaneously substituting b for a, this way obtaining $b \Vdash^{\forall} A$; similarly for world label substitution.

It is now routine to show that G3IL* enjoys height preserving substitution for world and neighbourhood labels.

PROPOSITION 5.2. *The following hold*:

(i) *If* $\overset{n}{\vdash} \Gamma \Rightarrow \Delta$, *then* $\overset{n}{\vdash} \Gamma[y/x] \Rightarrow \Delta[y/x]$;
(ii) *If* $\overset{n}{\vdash} \Gamma \Rightarrow \Delta$, *then* $\overset{n}{\vdash} \Gamma[b/a] \Rightarrow \Delta[b/a]$.

Proof. Straightforward induction on n. If $n = 0$, then $\Gamma \Rightarrow \Delta$ is an initial sequent, or a conclusion of $\mathcal{L}\bot$, $\mathcal{L}\varnothing$, or *Irrefl*. The same is true for $\Gamma[y/x] \Rightarrow \Delta[y/x]$ and for $\Gamma[b/a] \Rightarrow \Delta[b/a]$.[24] If $n > 0$, we consider the last rule applied. If the latter has no variable conditions, we apply the inductive hypothesis to the premise(s), followed by that very rule. Otherwise, the rule needs some care in case the substituted variable coincides with the fresh variable of the premise. In that case, we need to apply twice the inductive hypothesis to the premise, first to replace the fresh variable with another

[24] It is important to notice that $\varnothing$ is a *constant* of our language, and therefore it cannot be subject to substitution.

fresh variable—different from the one we wish to substitute—and secondly, to perform the desired substitution. □

5.1. General initial sequents, weakening, contraction, invertibility. Let φ denote either a relational atom or a (proper) labelled formula for the remaining sections.

We start with a relatively simple result.

LEMMA 5.3. *The following sequents are derivable in* $\mathsf{G3IL}^\star$:

1. $a \Vdash^\forall A, \Gamma \Rightarrow \Delta, a \Vdash^\forall A$;
2. $x : A, \Gamma \Rightarrow \Delta, x : A$.

Proof. The two cases are proven by mutual induction on the weight of the labelled formulas. The general strategy is to apply the left and right rule to treat the two formula occurrences until two formula occurrences of smaller weight are reached.

By means of example, we prove case 2, subcase $x : \langle]_i A$:

$$\dfrac{\dfrac{\begin{matrix}\vdots\ \text{IH}\\ xS_i a, a \Vdash^\forall A, \Gamma \Rightarrow \Delta, x : \langle]_i A, a \Vdash^\forall A\end{matrix}}{xS_i a, a \Vdash^\forall A, \Gamma \Rightarrow \Delta, x : \langle]_i A}\ \mathcal{R}\langle],}{x : \langle]_i A, \Gamma \Rightarrow \Delta, x : \langle]_i A}\ \mathcal{L}\langle]$$

where we can apply the inductive hypothesis to the top sequent since $\mathfrak{w}(a \Vdash^\forall A) < \mathfrak{w}(x : \langle]_i A)$.

Notice that the subcase $x : A \rhd_i B$ is easily managed since sequents $x : A \rhd_i B, \Gamma \Rightarrow \Delta, x : A \rhd_i B$ are initial, and hence derivable by design. □

We want now to establish the admissibility of weakening in $\mathsf{G3IL}^\star$.

LEMMA 5.4. *The rules of weakening*

$$\dfrac{\Gamma \Rightarrow \Delta}{\varphi, \Gamma \Rightarrow \Delta}\ \mathcal{L}Wk \qquad\qquad \dfrac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \varphi}\ \mathcal{R}Wk$$

are height-preserving admissible in $\mathsf{G3IL}^\star$.

Proof. We need to show that if $\overset{n}{\vdash} \Gamma \Rightarrow \Delta$, then $\overset{n}{\vdash} \varphi, \Gamma \Rightarrow \Delta$ and $\overset{n}{\vdash} \Gamma \Rightarrow \Delta, \varphi$. The proof consists of a straightforward induction on n, following the lines of the analogous results in [31, 33], which the reader is referred to for the details. □

Next, we can prove that all the rules of $\mathsf{G3IL}^\star$ are invertible.[25]

LEMMA 5.5. *All the rules of* $\mathsf{G3IL}^\star$ *are invertible.*

Proof. We proceed by induction on the height of the derivation, distinguishing cases based on the rule under consideration. Notice first that the rules for extensions

25 Recall from, e.g., [29], that a rule is invertible when, if its conclusion is derivable, so are its premise(s). This is a key feature of G3-style sequent calculi, whose main consequence is the dismissal of backtracking in a root-first proof search, as discussed by Troelstra and Schwichtenberg [49].

are clearly (height-preserving) invertible by (height-preserving) admissibility of weakening; the same remark applies to the rules for GVS, as well as to $\mathcal{L}\rhd$, $\mathcal{L}\Vdash^{\forall}$, and $\mathcal{R}\langle]$. Propositional cases are dealt with as in [28]. The cases $\mathcal{L}\langle]$ and $\mathcal{R}\Vdash^{\forall}$ are the same as in [31]. We can thus focus our proof on the invertibility of $\mathcal{R}\rhd$. Assume then that $\overset{n}{\vdash} \Gamma \Rightarrow \Delta, x : A \rhd_i B$. If $n = 0$ and $x : A \rhd_i B$ is not principal, then also $\overset{0}{\vdash} y \in R[x], y : A, y : A \rhd_i B, \Gamma \Rightarrow \Delta, y : \langle]_i B$. If it is principal, then $\Gamma = \Gamma', x : A \rhd_i B$ and we need to prove that

$$y \in R[x], y : A, y : A \rhd_i B, \Gamma', x : A \rhd_i B \Rightarrow \Delta, y : \langle]_i B$$

is derivable. But this is provable by application of $\mathcal{L}\rhd$ to the initial sequent

$$y \in R[x], y : A, y : A \rhd_i B, \Gamma', x : A \rhd_i B \Rightarrow \Delta, y : \langle]_i B, y : A \rhd_i B$$

with the sequents

$$y \in R[x], \mathbf{y : A}, y : A \rhd_i B, \Gamma', x : A \rhd_i B \Rightarrow \Delta, y : \langle]_i B, \mathbf{y : A}$$

and

$$\mathbf{y : \langle]_i B}, y \in R[x], y : A, y : A \rhd_i B, \Gamma', x : A \rhd_i B \Rightarrow \Delta, \mathbf{y : \langle]_i B},$$

which are both derivable in virtue of the highlighted labelled formulas occurring on both sides of each sequent (Lemma 5.3.(1))

If $n > 0$ and $x : A \rhd_i B$ is principal in the last rule, then we have the desired result. Otherwise, applying the inductive hypothesis to the premise(s) suffices. □

REMARK 3. *Notice that Lemma 5.5 cannot be strengthened into a height-preserving invertibility of the rules just because of the case we discussed in its proof: This is analogous to what happens for* G3GL *in* [28], *which we used as a model for the design of* G3IL*.

We now want to prove the admissibility of contraction. Before we proceed with the proof, it is appropriate to introduce a notion that will also be used in the proof of cut elimination.

DEFINITION 5.6 (After Negri and von Plato [33]). *The range $\mathfrak{r}(x)$ of a world label x in a derivation $\mathcal{D}$ in* G3IL* *is the set of world labels* y *such that either $y \in R[x]$ or for some $n \geq 1$ and for some $x_1, \ldots, x_n$ the relational atoms $x_1 \in R[x]$, $x_2 \in R[x_1]$,…, $y \in R[x_n]$ appear in the antecedent of sequents of $\mathcal{D}$. The range $\mathfrak{r}(a)$ of a neighbourhood label a in $\mathcal{D}$ is defined as $\mathfrak{r}(a) := \max\{\mathfrak{r}(x) \mid x \in a\} \cup \{*\}$ ordered w.r.t. set inclusion. We set $\mathfrak{r}(\{x\}) = \mathfrak{r}(x)$ and $\mathfrak{r}(\varnothing) = \varnothing$.*

Finally, we say that a rule is range-preserving admissible if the rule's elimination does not increase the ranges of labels in the derivation.

THEOREM 5.7. *The rules of contraction*

$$\frac{\varphi, \varphi, \Gamma \Rightarrow \Delta}{\varphi, \Gamma \Rightarrow \Delta}\ \mathcal{L}Ctr \qquad\qquad \frac{\Gamma \Rightarrow \Delta, \varphi, \varphi}{\Gamma \Rightarrow \Delta, \varphi}\ \mathcal{R}Ctr$$

are range-preserving admissible in G3IL*.

Proof. By simultaneous induction for left and right contraction, with primary induction on the weight of the contracted formula and secondary induction on the height of the derivation. The only case requiring some care is when the contracted formula is $x : A \rhd_i B$ in the consequent, of weight n. If the contracted formula is not principal for the last rule in the derivation—i.e. the latter is not $\mathcal{R}\rhd$ on $x : A \rhd_i B$—then we apply the secondary inductive hypothesis to the premise(s), followed by the rule. Otherwise, we invert the premise of $\mathcal{R}\rhd$ to obtain

$$y \in R[x], y : A, y : A \rhd_i B, y \in R[x], y : A, y : A \rhd_i B, \Gamma \Rightarrow \Delta, y : \langle]_i B, y : \langle]_i B.$$

By applying the main inductive hypothesis once for $y \in R[x]$, once for $y : A$, and once for $y : \langle]_i B$ to get

$$y : A \rhd_i B, y \in R[x], y : A, y : A \rhd_i B, \Gamma \Rightarrow \Delta, y : \langle]_i B.$$

Now, we claim that the following rule is height- and range-preserving admissible:

$$\frac{w : A \rhd_i B, w : A \rhd_i B, \Gamma \Rightarrow \Delta}{w : A \rhd_i B, \Gamma \Rightarrow \Delta}$$

and we prove it by primary induction on the weight of the contracted formula and secondary induction on the height of the derivation: If none instance of $w : A \rhd_i B$ is the main formula in the last rule of the derivation, then the height of the derivation is 0 and we can contract any formulas without increasing the height, as initial sequents are arbitrary weakened by design; if, on the contrary, one instance of $w : A \rhd_i B$ is the main formula of the last rule, then the derivation is as follows:

$$\frac{v \in R[w], w : A \rhd_i B, w : A \rhd_i B, \Gamma \Rightarrow \Delta, v : A \qquad v : \langle]_i B, v \in R[w], w : A \rhd_i B, w : A \rhd_i B, \Gamma \Rightarrow \Delta \qquad S}{v \in R[w], w : A \rhd_i B, w : A \rhd_i B, \Gamma \Rightarrow \Delta}\,\mathcal{L}\rhd_i$$

where S is the sequent $v \in R[w], w : A \rhd_i B, w : A \rhd_i B, \Gamma \Rightarrow \Delta, v : A \rhd_i B$.

We can apply the secondary induction hypothesis to each premise and then apply the $\mathcal{L}\rhd_i$ rule to get the desired conclusion.

Thus, we can use the just proven claim on

$$y \in R[x], y : A \rhd_i B, y \in R[x], y : A, y : A \rhd_i B, \Gamma \Rightarrow \Delta, y : \langle]_i B.$$

to derive then the conclusion of $\mathcal{R}Ctr$ by application of $\mathcal{R}\rhd_i$.

The range is preserved since, in inverting $\mathcal{R}\rhd$ and in the proof of the intermediate claim, we use labels already present in the derivation tree.

For the other cases, refer to e.g., [31, 33]. □

5.2. Cut elimination theorem. We have finally collected all the material required to prove the paper's main result, namely, cut elimination for $\mathsf{G3IL}^\star$.

THEOREM 5.8 (Cut admissibility). *The rule of cut*

$$\frac{\Gamma \Rightarrow \Delta, \varphi \qquad \varphi, \Gamma' \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,Cut$$

is admissible in $\mathsf{G3IL}^\star$.

Proof. The proof proceeds by primary induction on the weight of the cut formula, secondary induction on the range of the label of the cut formula,[26] and tertiary induction on the sum of heights of the cut premises.

Notice first that if y is in the range of x, then $\mathfrak{r}(y) \subsetneq \mathfrak{r}(x)$ because of the proof system design.

Following [32, 33], we start by distinguishing cases according to the rules applied to derive the premises of cut:

1. At least one of the premises of *Cut* is an initial sequent;
2. The cut formula is not principal in the derivation of at least one premise;
3. The cut formula is the principal formula of both derivations of the premises.

Case 1. Assume the leftmost premise of *Cut* is an initial sequent. If the cut formula is $x : p$ or $x : A \rhd_i B$, then by weakening the rightmost premise, we obtain the cut conclusion. If the sequent is initial in virtue of some labelled formula ψ occurring in both Γ and Δ, then the conclusion of *Cut* is an initial sequent too.

A similar argument works if we assume that the rightmost premise of *Cut* is an initial sequent.

If $x : \bot$ is the cut formula φ and the leftmost premise of *Cut* is not initial, we have derived it by some rule R. If R is $\mathcal{L}\bot$, then $x : \bot$ occurs in the conclusion of the cut, and therefore we can obtain that sequent from $\mathcal{L}\bot$. Similarly if R is $\mathcal{L}\varnothing$ or *Irrefl*. Otherwise, if R is different from $\mathcal{L}\bot$, $\mathcal{L}\varnothing$ and *Irrefl*, we can permute the cut up on the left premise and eliminate it by inductive hypotheses.

Case 2. Assume the cut formula is not principal in the last rule leading to the leftmost premise of *Cut*. The general situation is the following:

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\mathcal{D}_1\\ \Gamma^* \Rightarrow \Delta^*, \varphi\end{matrix}}{\Gamma \Rightarrow \Delta, \varphi}\,R \qquad \begin{matrix}\vdots\,\mathcal{D}_2\\ \varphi, \Gamma' \Rightarrow \Delta'\end{matrix}}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,Cut.$$

Then, we perform the following lifting of the cut and rely on the inductive hypotheses:

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\mathcal{D}_1\\ \Gamma^* \Rightarrow \Delta^*, \varphi\end{matrix} \qquad \begin{matrix}\vdots\,\mathcal{D}_2\\ \varphi, \Gamma' \Rightarrow \Delta'\end{matrix}}{\Gamma^*, \Gamma' \Rightarrow \Delta^*, \Delta'}\,Cut}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,R$$

[26] Recall, from Definition 5.6, that the range of a label is a set of labels occurring in atoms in the derivations. It is possible to reason by induction on the range w.r.t. set inclusion. Equivalently, we might consider the cardinality of the set identified by the range of a label and reason by induction w.r.t. the $\leq$-relation on natural numbers.

Next, we have the subcases of $Repl_1$ and $Repl_2$. We describe the situation with the right premise of *Cut* obtained by $Repl_1$ since the general setting follows the same line of reasoning. Assume then we have

$$\dfrac{\begin{matrix}\vdots\,\mathcal{D}_1\\ \Gamma \Rightarrow \Delta, x : p\end{matrix} \qquad \dfrac{\begin{matrix}\vdots\,\mathcal{D}_2\\ y \in \{x\}, x : p, y : p, \Gamma' \Rightarrow \Delta'\end{matrix}}{y \in \{x\}, x : p, \Gamma' \Rightarrow \Delta'}\,Repl_1}{y \in \{x\}, \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,Cut$$

Now, if $\mathcal{D}_1$ is an initial sequent, or if $x : \bot$, $x \in \varnothing$, or $x \in R[x]$ occurs in Γ we can weaken the right premise of *Cut*, or recur to the appropriate 0-ary $\mathcal{L}$-rule. Otherwise, we must have derived the left premise of the cut by some rule, and $x : p$ cannot be its principal formula. Lifting the cut does the job.

Case 3. We omit the propositional (sub)cases, referring to [33, theorem 11.9].

Notice first that we do not need to consider the rules for GVS nor the rules for extensions of $\mathbb{IL}$ since, by design, relational atoms only occur in the antecedent of sequents.

For the local forcing rules, we have

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\mathcal{D}_1\\ y \in a, \Gamma \Rightarrow \Delta, y : A\end{matrix}}{\Gamma \Rightarrow \Delta, a \Vdash^{\forall} A}\,\mathcal{R}{\Vdash^{\forall}} \qquad \dfrac{\begin{matrix}\vdots\,\mathcal{D}_2\\ x : A, x \in a, a \Vdash^{\forall} A, \Gamma' \Rightarrow \Delta'\end{matrix}}{x \in a, a \Vdash^{\forall} A, \Gamma' \Rightarrow \Delta'}\,\mathcal{L}{\Vdash^{\forall}}}{\Gamma, x \in a, \Gamma' \Rightarrow \Delta, \Delta'}\,Cut$$

which is solved by

$$\dfrac{\begin{matrix}\vdots\,\mathcal{D}_1[x/y]\\ x \in a, \Gamma \Rightarrow \Delta, x : A\end{matrix} \qquad \dfrac{\Gamma \Rightarrow \Delta, a \Vdash^{\forall} A \qquad \begin{matrix}\vdots\,\mathcal{D}_2\\ x : A, x \in a, a \Vdash^{\forall} A, \Gamma' \Rightarrow \Delta'\end{matrix}}{x \in a, x : A, \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,Cut,}{x \in a, x \in a, \Gamma, \Gamma, \Gamma' \Rightarrow \Delta, \Delta', \Delta}\,Cut$$

where the upper cut is on derivations of smaller height and on a label of the same range; the lower one is on formulas of smaller weight. Application(s) of contraction to the conclusion gives the cut conclusion.

For the intermediate modality, we have

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\mathcal{D}_1\\ yS_x a, \Gamma \Rightarrow \Delta, y : \langle]_x A, a \Vdash^{\forall} A\end{matrix}}{yS_x a, \Gamma \Rightarrow \Delta, y : \langle]_x A}\,\mathcal{R}\langle] \qquad \dfrac{\begin{matrix}\vdots\,\mathcal{D}_2\\ yS_x b, b \Vdash^{\forall} A, \Gamma' \Rightarrow \Delta'\end{matrix}}{y : \langle]_x A, \Gamma' \Rightarrow \Delta'}\,\mathcal{L}\langle]}{yS_x a, \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,Cut$$

which is solved by

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\mathcal{D}_1\\ yS_xa, \Gamma \Rightarrow \Delta, y : \langle]_x A, a \Vdash^{\forall} A\end{matrix} \qquad y : \langle]_x A, \Gamma' \Rightarrow \Delta'}{yS_xa, \Gamma, \Gamma' \Rightarrow \Delta, \Delta', a \Vdash^{\forall} A}\,\textit{Cut} \qquad \begin{matrix}\vdots\,\mathcal{D}_2[a/b]\\ yS_xa, a \Vdash^{\forall} A, \Gamma' \Rightarrow \Delta'\end{matrix}}{yS_xa, yS_xa, \Gamma, \Gamma', \Gamma' \Rightarrow \Delta, \Delta', \Delta'}\,\textit{Cut},$$

where the upper cut is on derivations of smaller height, and the lower one is on formulas of smaller weight. Application(s) of contraction to the conclusion gives the cut conclusion.

Finally, we consider the case of rules for $\rhd$. The general setting is

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\mathcal{D}_1\\ z \in R[x], z : A, z : A \rhd_i B, \Gamma \Rightarrow \Delta, z : \langle]_i B\end{matrix}}{\Gamma \Rightarrow \Delta, x : A \rhd_i B}\,\mathcal{R}\rhd \qquad \dfrac{\mathcal{D}_2 \qquad \mathcal{D}_3 \qquad \mathcal{D}_4}{y \in R[x], x : A \rhd_i B, \Gamma' \Rightarrow \Delta'}\,\mathcal{L}\rhd}{y \in R[x], \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,\textit{Cut},$$

where

$\mathcal{D}_2$ is a derivation of $y \in R[x], x : A \rhd_i B, \Gamma' \Rightarrow \Delta', y : A$;
$\mathcal{D}_3$ is a derivation of $y : \langle]_i B, y \in R[x], x : A \rhd_i B, \Gamma' \Rightarrow \Delta'$; and
$\mathcal{D}_4$ is a derivation of $y \in R[x], x : A \rhd_i B, \Gamma' \Rightarrow \Delta', y : A \rhd_i B$.

Perform the following steps:

a.
$$\dfrac{\Gamma \Rightarrow \Delta, x : A \rhd_i B \qquad \begin{matrix}\vdots\,\mathcal{D}_2\\ y \in R[x], x : A \rhd_i B, \Gamma' \Rightarrow \Delta', y : A\end{matrix}}{y \in R[x], \Gamma, \Gamma' \Rightarrow \Delta, \Delta', y : A}\,\textit{Cut},$$

where the cut is on derivations of smaller height.

b.
$$\dfrac{\Gamma \Rightarrow \Delta, x : A \rhd_i B \qquad \begin{matrix}\vdots\,\mathcal{D}_3\\ y \in R[x], y : \langle]_i B, x : A \rhd_i B, \Gamma' \Rightarrow \Delta'\end{matrix}}{y \in R[x], y : \langle]_i B, \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,\textit{Cut},$$

where the cut is on derivations of smaller height.

c.
$$\dfrac{\Gamma \Rightarrow \Delta, x : A \rhd_i B \qquad \begin{matrix}\vdots\,\mathcal{D}_3\\ y \in R[x], x : A \rhd_i B, \Gamma' \Rightarrow \Delta', y : A \rhd_i B\end{matrix}}{y \in R[x], \Gamma, \Gamma' \Rightarrow \Delta, \Delta', y : A \rhd_i B}\,\textit{Cut},$$

where the cut is on derivations of smaller height.

d. Finally proceed as follows:

$$\dfrac{\overset{\vdots\, a.}{\cdots \Rightarrow \cdots, y : A} \qquad \dfrac{\dfrac{\overset{\vdots\, c.}{\cdots \Rightarrow \cdots, y : A \rhd_i B} \qquad \overset{\vdots\, \mathcal{D}_1[y/z]}{y \in R[x], y : A, y : A \rhd_i B, \Gamma \Rightarrow \Delta, y : \langle]_i B}}{y \in R[x]^2, \Gamma^2, \Gamma', y : A \Rightarrow \Delta^2, \Delta', y : \langle]_i B}\,Cut \qquad \overset{\vdots\, b.}{y : \langle]_i B, \cdots \Rightarrow, \cdots}}{y \in R[x]^3, \Gamma^3, \Gamma'^2, y : A \Rightarrow \Delta^3, \Delta'^2}\,Cut}{y \in R[x]^4, \Gamma^4, \Gamma'^3 \Rightarrow \Delta^4, \Delta'^3}\,Cut,$$

where the first cut from the top is on a label of a smaller range, and the second and the third are on formulas of smaller weight.[27] We obtain the conclusion of the original cut by applying steps of contraction to the conclusion of the derivation described at point *d*.

□

Now that we know that the cut rule is admissible in $\mathsf{G3IL}^\star$, we can prove the admissibility of generalised replacement rules.

LEMMA 5.9. *The rules* Repl$_1$ *and* Repl$_2$ *generalised to all formulas of the language are admissible in* $\mathsf{G3IL}^\star$.

Proof. By simultaneous induction on the weight of formulas. Since contraction and cut are admissible for our calculi, it is enough to prove that the sequent $y \in \{x\}, \varphi(x) \Rightarrow \varphi(y)$ is derivable for any labelled formula φ, and then apply a *Cut*, followed by contraction steps, to the premise of generalised *Repl*$_1$. For *Repl*$_2$, we reason symmetrically.

The proof then proceeds by a straightforward induction on the weight of $\varphi(x)$ on the lines of [32, lemma 6.5.2]. □

REMARK 4. *Raymond Smullyan famously wrote in* [44]:

> *"The real importance of cut-free proofs is not the elimination of cuts per se, but rather that such proofs obey the subformula principle."*

For classical propositional logic, cut admissibility guarantees that the subformula principle holds in $\mathsf{G3cp}$, *as documented in* [49]. *For labelled calculi, that principle needs to be relaxed. In its original version, the subformula property holds whenever any sequent occurring in a derivation of a given sequent* $\Gamma \Rightarrow \Delta$ *contains only subformulas of the formulas composing the latter. In our systems, the rules for* $\rhd$ *introduce the intermediate modality* $\langle]_i$ *by decomposing* $A \rhd_i B$ *into* A *and* $\langle]_i B$, *and the latter is not, in strict terms, a subformula of* $A \rhd_i B$. *However, its weight is defined to be smaller than that of interpretability formulas: Therefore, we might say that it is less complex and, in that flexible sense, loosely generated by a* $\rhd$*-formula. Moreover, it is not hard to prove a pureness condition on labels: In* $\mathsf{G3IL}^\star$, *any derivation contains either eigenvariables in rules with freshness condition or labels already present in the conclusion. In this sense, the family of calculi we have designed can be considered analytic, in line with the standard* $\mathsf{G3}$ *paradigm.*

[27] For reasons of space, we had to replace some unessential formulas in sequents with dots.

§6. Completeness. In this section, we prove that our family of labelled sequent calculi is sound and complete w.r.t. the semantic and axiomatic presentations of interpretability logics under investigation.

We give syntactic proofs of those results and then discuss some aspects concerning a direct proof of semantic completeness for G3IL*.

6.1. Syntactic completeness. For a start, we need to interpret the underlying language of the labelled formulas into $\mathcal{L}_\rhd$, in which the modality is not indexed. Therefore, we agree to read $x : A \rhd B$ as $x : A \rhd_x B$, unless otherwise stated.

We immediately have the following theorem.

THEOREM 6.1 (Completeness). *If a formula A is derivable in $\mathbb{IL}$ or any of its extensions from Definition 2.4, then there is a derivation of the sequent $\Rightarrow x : A$ in the calculus* G3IL* *for the corresponding logic.*

Proof. We rely on the equivalence between $\mathbb{IL}$ and $\mathbb{IL}_\rhd$ and prove that all the axiom schemas and inference rules of the latter are derivable or proven to be admissible in G3IL. For the extensions, we only need to prove the axiom schema corresponding to the specific semantic rule of the family of labelled systems.

We recall the axiomatisation of $\mathbb{IL}_\rhd$ and its extensions from Definition 4.2 and §2.1:

$\mathbb{IL}_\rhd$

- standard axiomatisation of $\mathbb{CPC}$
- schema IL2 : $A \rhd B \to B \rhd C \to A \rhd C$
- schema IL3 : $A \rhd C \to B \rhd C \to A \vee B \rhd C$
- schema IL-Löb: $A \rhd (A \wedge (A \rhd \bot))$
- $\rhd$Rule $\dfrac{A \to B}{A \rhd B}$

Extensions

- schema M : $A \rhd B \to A \wedge (\neg C \rhd \bot) \rhd B \wedge (\neg C \rhd \bot)$
- schema P : $A \rhd B \to \neg(A \rhd B) \rhd \bot$
- schema W : $A \rhd B \to A \rhd B \wedge (A \rhd \bot)$
- schema KM1 : $A \rhd \neg(\top \rhd \bot) \to \top \rhd \neg A$
- schema M$_0$ $A \rhd B \to \neg(A \rhd \bot) \wedge (\neg C \rhd \bot) \rhd B \wedge (\neg C \rhd \bot)$.

For propositional logic, the derivability of axiom schemas is straightforward; the admissibility of cut assures the admissibility of modus ponens.

Next notice that the simpler $\mathcal{L}$-rule for $\rhd$

$$\frac{y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta, y : A \qquad y : \langle]_i B, y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta}{y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta}\ \mathcal{L}\rhd^S$$

is admissible in our systems, for any sequent of the form

$$y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta, y : A \rhd_i B$$

is derivable in $\mathsf{G3IL}^*$.

The same holds for the rule

$$\dfrac{y \in R[x], y : A, \Gamma \Rightarrow \Delta, y : \langle]_i B}{\Gamma \Rightarrow \Delta, x : A \rhd_i B}\ \mathcal{R}\rhd^{S}{}_{(y!)}$$

Now, we proceed with the basic calculus:

IL2: The derivation is rendered in Figure 5.
IL3: The derivation is rendered in Figure 6.
IL-Löb: The derivation is rendered in Figure 7.
$\rhd$ **Rule:** The derivation is as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots \ \text{By assumption } [i/x]}{\Rightarrow i : A \to B}}{i : A \Rightarrow i : B}\ \textit{Inversion } \mathcal{R}{\to}}{yS_x\{y\}, y \in R[x], y : A \rhd_x B, y : A, i \in \{y\}, i : A \Rightarrow i : B}\ \mathcal{L}Wk}{yS_x\{y\}, y \in R[x], y : A \rhd_x B, y : A, i \in \{y\} \Rightarrow i : B}\ \textit{Repl}_1}{yS_x\{y\}, y \in R[x], y : A \rhd_x B, y : A, i \in \{y\} \Rightarrow i : B, y : \langle]_x B}\ \mathcal{R}Wk}{yS_x\{y\}, y \in R[x], y : A \rhd_x B, y : A \Rightarrow y : \langle]_x B, \{y\} \Vdash^{\forall} B}\ \mathcal{R}{\Vdash^{\forall}}}{yS_x\{y\}, y \in R[x], y : A \rhd_x B, y : A \Rightarrow y : \langle]_x B}\ \mathcal{R}\langle]}{y \in R[x], y : A \rhd_x B, y : A \Rightarrow y : \langle]_x B}\ \textit{Qrefl}}{\Rightarrow x : A \rhd_x B}\ \mathcal{R}\rhd.$$

The derivations of each principle in the corresponding extension of G3IL are given in Appendix A. □

6.2. Soundness. We must interpret relational atoms and labelled formulas in Verbrugge models to prove the converse direction. The calculi per se do not have a direct formula interpretation in the language of $\texttt{Form}^{(i)}_{\rhd}$ but we can define an interpretation for them in GVS: We need a function that interprets the labels in Verbrugge frames, thus connecting the syntactic elements of the calculus with the semantic notions of §3.2.

DEFINITION 6.2. *Let $\mathcal{M} = \langle W, R, \{S_x\}_{x \in W}, v\rangle$ be a Verbrugge model for $\mathbb{IL}$ or its extensions, $\mathcal{W}$ a set of world labels, and $\mathcal{A}$ a set of neighbourhood labels. A $\mathcal{WA}$-* interpretation *over $\mathcal{M}$ consists of a pair of functions (ρ, σ) such that*:

- *$\rho : \mathcal{W} \to W$ maps each $i \in \mathcal{W}$ into a world $\rho(i) \in W$;*
- *$\sigma : \mathcal{A} \to \wp(W)$ maps each $a \in \mathcal{A}$ into a nonempty set of worlds $\sigma(a) \in \wp(R[x])$, for $x \in W$, and $\varnothing$ into $\varnothing$.*

The notion of satisfiability *of a labelled formula under a $\mathcal{WA}$-interpretation is defined by cases on the form of that formula*:

$$
\begin{array}{cl}
yS_xb, zS_xb, \mathbf{b} \Vdash^{\forall} \mathbf{C}, z : B, z \in R[x], a \subseteq R[x], z \in a, yS_xa, a \Vdash^{\forall} B, y \in R[x]^2, y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, z : C, \mathbf{b} \Vdash^{\forall} \mathbf{C} & \\
\hline
yS_xb, zS_xb, b \Vdash^{\forall} C, z : B, z \in R[x], a \subseteq R[x], z \in a, yS_xa, a \Vdash^{\forall} B, y \in R[x]^2, y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, z : C & \mathcal{R}\langle] \\
\hline
zS_xb, b \Vdash^{\forall} C, z : B, z \in R[x], a \subseteq R[x], z \in a, yS_xa, a \Vdash^{\forall} B, y \in R[x]^2, y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, z : C & \mathit{Qtrans6} \\
\hline
z : \langle]_x C, z : B, z \in R[x], a \subseteq R[x], z \in a, yS_xa, a \Vdash^{\forall} B, y \in R[x]^2, y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, z : C \qquad z : B, \cdots \Rightarrow \cdots, z : B & \mathcal{L}\langle] \\
\hline
z : B, z \in R[x], a \subseteq R[x], z \in a, yS_xa, a \Vdash^{\forall} B, y \in R[x]^2, y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, z : C & \mathcal{L}\rhd^{S} \\
\hline
z \in R[x], a \subseteq R[x], z \in a, yS_xa, a \Vdash^{\forall} B, y \in R[x]^2, y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, z : C & \mathcal{L}\Vdash^{\forall} \\
\hline
a \subseteq R[x], z \in a, yS_xa, a \Vdash^{\forall} B, y \in R[x]^2, y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, z : C & \mathcal{L}\subseteq \\
\hline
z \in a, yS_xa, a \Vdash^{\forall} B, y \in R[x], y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, z : C & \mathit{DefS}_1 \\
\hline
yS_xa, a \Vdash^{\forall} B, y \in R[x], y : A, x : A \rhd_x B, x : B \rhd_x C \Rightarrow y : \langle]_x C, a \Vdash^{\forall} C & \mathcal{R}\Vdash^{\forall} \\
\hline
yS_xa, a \Vdash^{\forall} B, y \in R[x], y : A, x : A \rhd_x B, x : B_{\rhd} C \Rightarrow y : \langle]_x C & \mathcal{R}\langle] \\
\hline
y : \langle]_x B, y \in R[x], y : A, x : A \rhd_x B, x : B_{\rhd} C \Rightarrow y : \langle]_x C \qquad y : A, \cdots \Rightarrow \cdots, y : A & \mathcal{L}\langle] \\
\hline
y \in R[x], y : A, x : A \rhd_x B, x : B_{\rhd} C \Rightarrow y : \langle]_x C & \mathcal{L}\rhd^{S} \\
\hline
x : A \rhd_x B, x : B \rhd_x C \Rightarrow x : A \rhd_x C & \mathcal{R}\rhd^{S}
\end{array}
$$

Figure 5. *Derivation of* IL2.

$$
\dfrac{\dfrac{\dfrac{\dfrac{y : B, \cdots \Rightarrow \cdots, y : B \qquad y : A, \cdots \Rightarrow \cdots, y : B}{y \in R[x], y : A \vee B, x : A \rhd_x C, x : B \rhd_x C \Rightarrow y : \langle]_x C, y : A, y : B}\ \mathcal{L}\vee \qquad y : \langle]_x C, \cdots \Rightarrow y : \langle]_x C, y : A}{y \in R[x], y : A \vee B, x : A \rhd_x C, x : B \rhd_x C \Rightarrow y : \langle]_x C, y : A}\ \mathcal{L}\rhd^S \qquad y : \langle]_x C, y \in R[x], y : A \vee B, x : A \rhd_x C, x : B \rhd_x C \Rightarrow y : \langle]_x C}{y \in R[x], y : A \vee B, x : A \rhd_x C, x : B \rhd_x C \Rightarrow y : \langle]_x C}\ \mathcal{L}\rhd^S}{x : A \rhd_x C, x : B \rhd_x C \Rightarrow x : A \vee B \rhd_x C}\ \mathcal{R}\rhd^S
$$

Figure 6. *Derivation of* IL3.

$$yS_xb, yS_x\{z\}, z \in \{z\}, zS_xb, \mathbf{b} \Vdash^{\forall} (\mathbf{A} \wedge (\mathbf{A} \rhd \bot)), z \in R[y], z \in R[i], z : A, i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), z : \langle]_i\bot, \mathbf{b} \Vdash^{\forall} (\mathbf{A} \wedge (\mathbf{A} \rhd \bot))$$

$\mathcal{R}\langle]$

$$yS_xb, yS_x\{z\}, z \in \{z\}, zS_xb, b \Vdash^{\forall} (A \wedge (A \rhd \bot)), z \in R[y], z \in R[i], z : A, i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), z : \langle]_i\bot$$

$Qtrans6$

$$yS_x\{z\}, z \in \{z\}, zS_xb, b \Vdash^{\forall} (A \wedge (A \rhd \bot)), z \in R[y], z \in R[i], z : A, i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), z : \langle]_i\bot$$

$DefS_2$

$$z \in \{z\}, zS_xb, b \Vdash^{\forall} (A \wedge (A \rhd \bot)), z \in R[y], z \in R[i], z : A, i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), z : \langle]_i\bot$$

$Sing$

$$zS_xb, b \Vdash^{\forall} (A \wedge (A \rhd \bot)), z \in R[y], z \in R[i], z : A, i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), z : \langle]_i\bot$$

$\mathcal{R}\langle]$

$$z : \langle]_x(A \wedge (A \rhd \bot)), z \in R[y], z \in R[i], z : A, i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), z : \langle]_i\bot \qquad z : A, \cdots \Rightarrow \cdots, z : A$$

$\mathcal{L}\rhd^S$

$$z \in R[y], z \in R[i], z : A, i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), z : \langle]_i\bot$$

$Repl_2$

$$z \in R[i], z : A, i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), z : \langle]_i\bot$$

$\mathcal{R}\rhd^S$

$$i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), i : A \rhd_i \bot$$

Right premise:

$$i : A, \cdots \Rightarrow \cdots, i : A$$

$Repl_2$

$$i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), i : A$$

$\mathcal{R}\wedge$

$$i \in \{y\}, yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), i : (A \wedge (A \rhd \bot))$$

$\mathcal{R}\Vdash^{\forall}$

$$yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot)), \{y\} \Vdash^{\forall} (A \wedge (A \rhd \bot))$$

$\mathcal{R}\langle]$

$$yS_x\{y\}, y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot))$$

$Qrefl$

$$y \in R[x], y : A, y : A \rhd_x (A \wedge (A \rhd \bot)) \Rightarrow y : \langle]_x(A \wedge (A \rhd \bot))$$

$\mathcal{R}\rhd$

$$\Rightarrow x : A \rhd_x (A \wedge (A \rhd \bot))$$

Figure 7. *Derivation of* IL-*Löb*.

- $\rho, \sigma \vDash_{\mathcal{M}} y \in R[x]$ *if* $\rho(x) R \rho(y)$;
- $\rho, \sigma \vDash_{\mathcal{M}} y \in a$ *if* $\rho(y) \in \sigma(a)$;
- $\rho, \sigma \vDash_{\mathcal{M}} y S_x a$ *iff* $y \in S_x^{-1} a$: *if* $\rho(y) S_{\rho(x)} \sigma(a)$;
- $\rho, \sigma \vDash_{\mathcal{M}} y \in R[a]$ *if, for some* $x \in \sigma(a)$, $x R \rho(y)$;
- $\rho, \sigma \vDash_{\mathcal{M}} a \subseteq b$ *if* $\sigma(a) \subseteq \sigma(b)$;
- $\rho, \sigma \vDash_{\mathcal{M}} y \in \{x\}$ *if* $\rho(y) \in \{\rho(x)\}$;
- $\rho, \sigma \vDash_{\mathcal{M}} y S_i \{x\}$ *if* $\rho(y) S_{\rho(i)} \{\rho(x)\}$;
- $\rho, \sigma \vDash_{\mathcal{M}} x : p$ *if* $\rho(x) \Vdash_{\mathcal{M}} p$, *and similarly for formulas obtained by classical propositional connectives*;
- $\rho, \sigma \vDash_{\mathcal{M}} a \Vdash^{\forall} A$ *if* $\rho(a) \Vdash^{\forall} A$;
- $\rho, \sigma \vDash_{\mathcal{M}} x : \langle]_i A$ *if, for some* $a \in (\wp(R[\rho(i)]) \setminus \{\varnothing\})$, $\rho(x) S_{\rho(i)} a$ *and* $a \Vdash^{\forall} A$;
- $\rho, \sigma \vDash_{\mathcal{M}} x : A \rhd_i B$ *if, for all* $y \in R[\rho(x)]$, *if* $y \Vdash_{\mathcal{M}} A$, *then* $y \Vdash_{\mathcal{M}} \langle]_{\rho(i)} B$.

Given a sequent $\Gamma \Rightarrow \Delta$, let $\mathcal{W}, \mathcal{A}$ be the sets of world and neighbourhood labels occurring in $\Gamma \cup \Delta$, and let (ρ, σ) be a $\mathcal{WA}$-interpretation.

Define $\rho, \sigma \vDash_{\mathcal{M}} \Gamma \Rightarrow \Delta$ if either $\rho, \sigma \nvDash_{\mathcal{M}} \varphi$ for some $\varphi \in \Gamma$ or $\rho, \sigma \vDash_{\mathcal{M}} \psi$ for some $\psi \in \Delta$. Define then validity under all interpretations by $\vDash_{\mathcal{M}} \Gamma \Rightarrow \Delta$ if and only if $\rho, \sigma \vDash_{\mathcal{M}} \Gamma \Rightarrow \Delta$ for all (ρ, σ). Finally, let us say that a sequent is valid in all Verbrugge models if $\vDash_{\mathcal{M}} \Gamma \Rightarrow \Delta$ for all models based on GVS appropriate to a specific interpretability logic.

THEOREM 6.3 (Soundness). *If a sequent* $\Rightarrow x : A$ *is derivable in* $\mathsf{G3IL}^{\star}$, *then* A *is a theorem of* $\mathbb{IL}$ *or of the corresponding extension*.

Proof. We prove something stronger, namely:

> (♠) If a sequent $\Gamma \Rightarrow \Delta$ is derivable in $\mathsf{G3IL}^{\star}$, then it is valid in the corresponding class of Verbrugge models.

The main theorem follows by modal completeness of the axiomatic calculi for interpretability we are dealing with.

Now, the proof of (♠) is by straightforward induction on the height of the derivation, by recurring to the notion of interpretation defined before. By means of example, we show the soundness of the left and right rule for the $\rhd$-operator.

$\mathcal{L}\rhd$ By inductive hypothesis we have

$$\rho, \sigma \vDash_{\mathcal{M}} y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta, y : A;$$

$$\rho, \sigma \vDash_{\mathcal{M}} y : \langle]_i B, y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta;$$

$$\rho, \sigma \vDash_{\mathcal{M}} y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta, y : A \rhd_i B.$$

There are two relevant cases to consider.

One has $\rho, \sigma \vDash_{\mathcal{M}} y : A$, $\rho, \sigma \vDash_{\mathcal{M}} y \in R[x]$ and $\rho, \sigma \nvDash_{\mathcal{M}} \langle]_i B$. From the former we have that $\rho(y) \Vdash A$; from the latter that $\rho(y) \nVdash \langle]_i B$. By definition, this means that $\rho, \sigma \nvDash_{\mathcal{M}} x : A \rhd_i B$, so that $\rho, \sigma \vDash_{\mathcal{M}} y \in R[x], x : A \rhd_i B, \Gamma \Rightarrow \Delta$.

The second one is $\rho, \sigma \vDash_{\mathcal{M}} y \in R[x]$ and $\rho, \sigma \nvDash_{\mathcal{M}} y : A \rhd_i B$. From the latter we have that $\rho(y) \nVdash A \rhd_i B$, from which we know that there exists a $z \in R[y]$ such that $z \Vdash A$ and $z \nVdash \langle]_i B$. By transitivity, we can reduce the situation to the first one.

$\mathcal{R}\rhd$ Assume by inductive hypothesis that $\vDash_{\mathcal{M}} y \in R[x], y : A \rhd_i B, y : A, \Gamma \Rightarrow \Delta, y : \langle]_i B$. We reason by contradiction. If none of Δ and $x : A \rhd_i B$ is valid under this interpretation, then there exists a $w \in R[\rho(x)]$ such that $w \Vdash A$ and $w \nVdash \langle]_i B$. Then we can define a new ρ' which is identical to ρ except possibly on y, for which we set $\rho'(y) = w$. However, by assumption on validity of the premise of the rule, since $w \Vdash A$ and $w \nVdash \langle]_i B$, we have $w \nVdash A \rhd_i B$. It is now straightforward to see that an ascending R-chain can be built on the model, contrary to the assumption that $\mathcal{M}$ is finite—or, alternatively, Noetherian.

□

REMARK 5. *The literature confirms that* $\mathbb{IL}$ *is sound and complete with respect to the version of Verbrugge semantics based on the quasi-transitivity condition* 6 *that we have adopted. Consequently, our system* G3IL *is sound with respect to* $\mathbb{IL}$*. However, we cannot extend this claim to the other axiomatic extensions considered here, as the existing proofs of completeness for those logics specifically address the version of Verbrugge semantics based on condition* 2 *for quasi-transitivity. Notably, a margin note in* [50] *indicates that for* $\mathbb{ILKM}1$*, GVS based on condition* 6 *cannot distinguish this logic from* $\mathbb{ILM}$.[28] *Thus, a key focus of our future work will be to incorporate the more expressive condition for quasi-transitivity within our framework.*

6.3. On termination. Labelled sequent calculi have a peculiar characteristic: From a failed proof search, extracting a countermodel to the sequent at the root of the derivation tree is generally possible. This way, a direct and constructive proof of modal completeness is obtained.[29]

For our family G3IL*, we could claim that the same holds: It is possible to define a Tait–Schütte–Takeuti procedure to construct a refutation of a sequent from a failed proof-search; however, in the present setting, it is essential to define that saturation process along with a terminating strategy for performing a proof search in G3IL*.

The reader familiar with this kind of task will see at once that our systems are rather complex to handle from a combinatorial perspective, and this imposes several cases to check when proving that root-first proof search in G3IL* does terminate.[30]

It is essential to present an equivalent basic proof system for IL, that—following [28, sec. 5]—one might call G3KIL. By replacing $\mathcal{L}\rhd$ with the simplified rule $\mathcal{L}\rhd^S$ that we used in the proof of Theorem 6.1, we get the basic calculus; extensions are built on top of it, according to the rules in Figure 4.

We are confident that G3KIL—as well as its extensions G3KIL$^{\mathsf{M}}$ and G3KIL$^{\mathsf{P}}$—can be shown to satisfy terminating proof search by adapting the proof strategy of Girlando et al. [10] to our much more intricate setting. A complete proof of termination for those systems should be easily extended to the other interpretability logic constituting our uniform family. A further step in difficulty would probably materialise when dealing with the labelled sequent calculus for $\mathbb{ILP}_0$, since—in order to define the generalised frame condition corresponding to the schema P_0 utilising a geometric rule—we would

28 We thank the reviewers for highlighting this explicit example.
29 Refer [30] for an extensive discussion.
30 In particular, it is rather difficult to handle the interaction of labels because of the presence of the indexed relation S, as well as the need to label the interpretability modality to express Noetherian induction by the $\mathcal{R}\rhd$ rule.

need further rules for set-theoretic operations on world sets, and thus further cases of interaction between world and neighbourhood labels.

Since, at present, we can only provide an informal argument supporting our claims, in complete earnestness, we prefer to propose the termination of proof search in our family of labelled sequent calculi as the following.

CONJECTURE 6.4 (Termination). *There exists a strategy making proof search in* $\mathsf{G3KIL}^\star$ *for a sequent of the form* $\Rightarrow x : A$ *always terminate in a finite number of steps. Moreover, it is possible to extract a countermodel to A belonging to the appropriate class of Verbrugge frames from a failed proof search.*

§7. Conclusions. In the present work, we have introduced a family $\mathsf{G3IL}^\star$ of labelled sequent calculi for interpretability logics based on Verbrugge semantics of [50], which subsumes the standard relational semantics for those modal logics discussed in [7].

Our systems are modularly obtained from a basic calculus G3IL according to the methodology of [32, chap. 6] for formalising geometric theories in the G3 paradigm for sequent calculi. We have proved that all these calculi do satisfy the main desiderata for good G3-sequent systems; in particular, we have given a detailed, syntactic proof of the cut-admissibility $\mathsf{G3IL}^\star$ (Theorem 5.8), for which we used a ternary measure of complexity of the cut instance. Our proof is constructive and it thus provides a cut-elimination algorithm for equivalent formulations of these labelled sequent calculi including a cut rule: We have an effective procedure for removing any instance of the cut rule from formal derivations in these (apparently richer) systems which can be, in principle, be implemented in as a proper computer algorithm.[31]

To our knowledge, there are no proof-theoretic treatments of interpretability logics as extensive as the one we have presented here.

What we have discussed in this paper can also be viewed as a case study for exploring the potential of explicit internalization of semantics in sequent calculi. Indeed, it is not immediately clear how to design a sequent system for the logics we have considered using the methodology of implicit internalization. To the best of our knowledge, no attempt has yet been made in the literature to achieve this.

7.1. Related work. The preprint [15] defines a labelled system for IL and some extensions as prefixed tableaux that internalise standard Veltman semantics. Our approach can be viewed as its direct dual—due to the duality between labelled tableaux and labelled sequent calculi—but it differs in its focus on the structural properties of the calculi and is based on GVS. With the adjustments discussed below, this approach allows us to cover a broader range of interpretability principles, notably the system $\mathbb{ILW}$ based on the de Jongh–Visser schema, which cannot be characterised by a first-order formula in the language of standard Veltman semantics.

On a more traditional note, Sasaki [39] presents a Gentzen-style sequent calculus for the basic interpretability logic IL. This work does not consider any extensions, and, as is typical for standard sequent calculi, it is not straightforward to modularly extend the calculus discussed in that paper to encompass additional interpretability logics.[32]

---

31 Refer to the standard treatment of the relation between cut-elimination and cut-admissibility given by Troelstra and Schwichtenberg [49, chap. 4].

32 The paper [1] discusses Craig interpolation and fixed-point properties for some interpretability logics; similarly, the recent work [16] addresses the same properties for subsystems of

7.2. Future work. In the near future, we would be pleased to resolve Conjecture 6.4 by providing a comprehensive proof of the termination of proof search in our sequent calculi. Settling this conjecture would not merely be a technical achievement; although the decidability of all the interpretability logics under consideration is known, their analysis from the perspective of complexity theory remains a work in progress, as demonstrated by the work of Bou and Joosten [6], Mikec et al. [22] and Mikec [25].

Nevertheless, as we noted above, *even prior* to resolving the termination conjecture, it is more crucial to extend our proof systems for interpretability logics to formalise, through labelled sequent calculi, the most expressive version of Verbrugge semantics, based on the quasi-transitivity condition 2 discussed in §4.1. This semantics is undoubtedly capable of distinguishing the interpretability principles explored in the most advanced literature on the topic.[33]

Furthermore, it may be pertinent to apply the labelling technique we have adopted here to provide a constructive proof of modal completeness for interpretability logics. This area is indeed a flourishing research field, and several model-theoretic techniques have been developed in recent years, as summarised by Joosten et al. [17]. Despite progressively achieving a modular character, parts of these proofs remain somewhat obscure and sensitive to the logic under investigation, as they rely on (variations of) Henkin's method.[34] A "reverse engineering" approach, relying on the design methodology employed in this work, may prove beneficial in developing alternative modular proofs of completeness and even in addressing some open problems related to model-theoretic matters for specific interpretability principles. In particular, it would be intriguing to test the capabilities of our methodology for proof system design concerning more exotic interpretability principles, such as those in the so-called broad series by Goris and Joosten [13], or the subsystems of IL discussed by [19].

Finally, from a more applicative perspective, it would be interesting to consider implementations of automated theorem provers for interpretability logics based on the mechanisation of our calculi, following the lines of previous work by [4, 20, 21] within a general-purpose proof assistant.

§A. Appendix: Proofs of syntactic completeness for extensions.

We present here the formal derivations of the interpretability principles for the extensions considered in Definition 2.4. For space reasons, we have to omit some components of some sequents in order to make the derivations fit the page layout. Notice thus that the best way to read the following proof trees is root-first, in a bottom-up construction. We highlighted in boldface the formulas in wide sequents closing derivation branches.

A.1. G3ILM. The derivation in G3ILM of the schema M is given in Figure A1.

A.2. G3ILP. The derivation in G3ILP of the schema P is given in Figure A2.

IL. Both papers employ model-theoretic arguments and techniques to achieve these proof-theoretic results.

33 We do not exclude the possibility of refining our systems to capture the more expressive quasi-transitivity condition (*Qtrans2*) by adapting (variants of) the more sophisticated sequent enrichment explored by Perini Brogi et al. [36].

34 Joosten and collaborators based their proofs on a 'labelling technique' that bears no relation to labelled sequent calculi; it is merely coincidental that both their approach and ours involve labels for interpretability logics.

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\mathcal{D}_1
\qquad
\dfrac{
\dfrac{\mathbf{z} : \mathbf{B}, z \in a, z \in b, \cdots, b \subseteq a, \cdots, a \Vdash^{\forall} B, \cdots \Rightarrow \cdots, \mathbf{z} : \mathbf{B}}{z \in a, z \in b, \cdots, b \subseteq a, \cdots, a \Vdash^{\forall} B, \cdots \Rightarrow \cdots, z : B}\ \scriptstyle\mathcal{L}\Vdash^{\forall}
}{z \in b, \cdots, b \subseteq a, \cdots, a \Vdash^{\forall} B, \cdots \Rightarrow \cdots, z : B}\ \scriptstyle\mathcal{L}\subseteq
}{z \in b, yS_x b, b \subseteq a, R[b] \subseteq R[y], yS_x a, a \Vdash^{\forall} B, y \in R[x], y : A, y : (\neg C \rhd_y \bot), x : A \rhd_x B \Rightarrow y : \langle]_x B \wedge (\neg C \rhd \bot), z : B \wedge (\neg C \rhd \bot)}\ \scriptstyle\mathcal{R}\wedge
}{yS_x b, b \subseteq a, R[b] \subseteq R[y], yS_x a, a \Vdash^{\forall} B, y \in R[x], y : A, y : (\neg C \rhd_y \bot), x : A \rhd_x B \Rightarrow y : \langle]_x B \wedge (\neg C \rhd \bot), b \Vdash^{\forall} B \wedge (\neg C \rhd \bot)}\ \scriptstyle\mathcal{R}\Vdash^{\forall}
}{yS_x b, b \subseteq a, R[b] \subseteq R[y], yS_x a, a \Vdash^{\forall} B, y \in R[x], y : A, y : (\neg C \rhd_y \bot), x : A \rhd_x B \Rightarrow y : \langle]_x B \wedge (\neg C \rhd \bot)}\ \scriptstyle\mathcal{R}\langle]
}{yS_x a, a \Vdash^{\forall} B, y \in R[x], y : A, y : (\neg C \rhd_y \bot), x : A \rhd_x B \Rightarrow y : \langle]_x B \wedge (\neg C \rhd \bot)}\ M
}{y : \langle]_x B, y \in R[x], y : A, y : (\neg C \rhd_y \bot), x : A \rhd_x B \Rightarrow y : \langle]_x B \wedge (\neg C \rhd \bot)}\ \scriptstyle\mathcal{L}\langle]
\qquad
y \in R[x], \mathbf{y} : \mathbf{A}, y : (\neg C \rhd_y \bot), x : A \rhd_x B \Rightarrow y : \langle]_x B \wedge (\neg C \rhd \bot), \mathbf{y} : \mathbf{A}
}{
\dfrac{
\dfrac{y \in R[x], y : A, y : (\neg C \rhd_y \bot), x : A \rhd_x B \Rightarrow y : \langle]_x B \wedge (\neg C \rhd \bot)}{y \in R[x], y : A \wedge (\neg C \rhd \bot), x : A \rhd_x B \Rightarrow y : \langle]_x B \wedge (\neg C \rhd \bot)}\ \scriptstyle\mathcal{L}\wedge
}{x : A \rhd_x B \Rightarrow x : A \wedge (\neg C \rhd \bot) \rhd_x B \wedge (\neg C \rhd \bot)}\ \scriptstyle\mathcal{R}\rhd^{S}
}\ \scriptstyle\mathcal{L}\rhd^{S}
$$

Figure A1. *Derivation of* M.

$$\dfrac{\mathbf{i{:}B}, i \in b, b \subseteq a, zS_y b, zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B, \mathbf{i{:}\ B}}{i \in b, b \subseteq a, zS_y b, zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B, i : B} \ \mathcal{L}\subseteq, \mathcal{L}\Vdash^{\forall}$$

$$\dfrac{i \in b, b \subseteq a, zS_y b, zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B, i : B}{b \subseteq a, zS_y b, zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B, b \Vdash^{\forall} B} \ \mathcal{R}\Vdash^{\forall}$$

$$\dfrac{b \subseteq a, zS_y b, zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B, b \Vdash^{\forall} B}{b \subseteq a, zS_y b, zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B} \ \mathcal{R}\langle]$$

$$\dfrac{b \subseteq a, zS_y b, zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B}{zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B} \ P$$

$$\dfrac{zS_x a, a \Vdash^{\forall} B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B}{z : \langle]_x B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B} \ \mathcal{L}\langle]$$

$$\dfrac{z : \langle]_x B, z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B \qquad z \in R[y], \mathbf{z{:}A}, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B), \mathbf{z{:}A}}{z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B)} \ \mathcal{L}\rhd^{S}$$

$$\dfrac{z \in R[y], z : A, y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, z : \langle]_y B)}{y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, y : (A \rhd_y B)} \ \mathcal{L}\rhd^{S}$$

$$\dfrac{y \in R[x], x : A \rhd_x B \Rightarrow y : \langle]_x \bot, y : (A \rhd_y B)}{y \in R[x], y : \neg(A \rhd B), x : A \rhd_x B \Rightarrow y : \langle]_x \bot} \ \mathcal{L}\neg$$

$$\dfrac{y \in R[x], y : \neg(A \rhd B), x : A \rhd_x B \Rightarrow y : \langle]_x \bot}{x : A \rhd_x B \Rightarrow x : \neg(A \rhd B) \rhd_x \bot} \ \mathcal{R}\rhd^{S}$$

Figure A2. *Derivation of* P.

$$\dfrac{\mathbf{j} : \bot, j \in c, iS_y c, c \Vdash^{\forall} \bot, \cdots \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}{j \in c, iS_y c, c \Vdash^{\forall} \bot, \cdots \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}\ \mathcal{L}\Vdash^{\forall}$$

$$\dfrac{j \in c, iS_y c, c \Vdash^{\forall} \bot, \cdots \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}{iS_y c, c \Vdash^{\forall} \bot, \cdots \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}\ NE$$

$$\dfrac{iS_y c, c \Vdash^{\forall} \bot, \cdots \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}{i : \langle]_y \bot, \cdots \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}\ \mathcal{L}\langle]$$

$$\dfrac{}{\cdots \Rightarrow i : C, i : \neg C}\ \mathcal{R}\neg$$

$$\dfrac{i : \langle]_y \bot, \cdots \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot \qquad \cdots \Rightarrow i : C, i : \neg C}{i \in R[b], i \in R_y, i \in R[w], w \in b, b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}\ \mathcal{L}\rhd^{S}$$

$$\dfrac{i \in R[b], i \in R_y, i \in R[w], w \in b, b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}{i \in R[w], w \in b, b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}\ \mathit{Trans}, \mathcal{L}\subseteq$$

$$\dfrac{i \in R[w], w \in b, b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, i : C, i : \langle]_w \bot}{w \in b, b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, w : B \wedge (\neg C \rhd \bot)}\ \mathcal{R}\rhd^{S}, \mathcal{R}\neg$$

$$\dfrac{w \in b, b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, w : B \wedge (\neg C \rhd \bot) \qquad \mathbf{w{:}B}, \cdots \Rightarrow \cdots, \mathbf{w{:}B}}{w \in b, b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, w : B \wedge (\neg C \rhd \bot)}\ \mathcal{R}\wedge$$

$$\dfrac{w \in b, b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot, w : B \wedge (\neg C \rhd \bot)}{b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot}\ \mathcal{R}\langle], \mathcal{R}\Vdash^{\forall}$$

$$\dfrac{b \subseteq a, yS_x b, R[b] \subseteq R[y], zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot}{zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot}\ M_0$$

$$\dfrac{zS_x a, a \Vdash^{\forall} B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot}{z : \langle]_x B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot}\ \mathcal{L}\langle]$$

$$\dfrac{z : \langle]_x B, z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot \qquad S}{z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot}\ \mathcal{L}\rhd^{S}$$

$$\dfrac{z \in R[x], z \in R[y], z : A, y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), z : \langle]_y \bot}{y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), y : A \rhd_y \bot}\ \mathcal{R}\rhd^{S}, \mathit{Trans}$$

$$\dfrac{y \in R[x], y : \neg C \rhd_y \bot, x : A \rhd_x B \Rightarrow y : \langle]_x(B \wedge (\neg C \rhd \bot)), y : A \rhd_y \bot}{x : A \rhd_x B \Rightarrow x : \neg(A \rhd \bot) \wedge (\neg C \rhd \bot) \rhd_x (B \wedge (\neg C \rhd \bot))}\ \mathcal{R}\rhd^{S}, \mathcal{L}\neg$$

Figure A3. *Derivation of* M_0.

$$\mathbf{i}: \bot, i \in b, wS_y b, b \Vdash^{\forall} \bot, w \in R[y], w \in R[z], \cdots \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots, z: B \rhd_z \bot, w: \langle]_z \bot \quad \mathcal{L}\Vdash^{\forall}$$

$$i \in b, wS_y b, b \Vdash^{\forall} \bot, w \in R[y], w \in R[z], \cdots \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots, z: B \rhd_z \bot, w: \langle]_z \bot \quad NE$$

$$wS_y b, b \Vdash^{\forall} \bot, w \in R[y], w \in R[z], \cdots \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots, z: B \rhd_z \bot, w: \langle]_z \bot \quad \mathcal{L}\langle]$$

$$w: \langle]_y \bot, w \in R[y], w \in R[z], \cdots \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots, z: B \rhd_z \bot, w: \langle]_z \bot \qquad w \in R[y], \cdots, y: (B \rhd_y \bot), \cdots \Rightarrow \cdots, \mathbf{w{:}B} \quad \mathcal{L}\rhd^{S}$$

$$w \in R[y], w \in R[z], w: B, z \in a, R[z] \subseteq R[y], yS_x a, a \Vdash^{\forall} \neg(B \rhd \bot), y: (B \rhd_y \bot), y: A, y \in R[x], x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots, z: B \rhd_z \bot, w: \langle]_z \bot \quad \mathcal{L}\subseteq$$

$$w \in R[z], w: B, z \in a, R[z] \subseteq R[y], yS_x a, a \Vdash^{\forall} \neg(B \rhd \bot), y: (B \rhd_y \bot), y: A, y \in R[x], x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots, z: B \rhd_z \bot, w: \langle]_z \bot \quad \mathcal{R}\rhd^{S}$$

$$z \in a, R[z] \subseteq R[y], yS_x a, a \Vdash^{\forall} \neg(B \rhd \bot), y: (B \rhd_y \bot), y: A, y \in R[x], x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots, z: B \rhd_z \bot \quad \mathcal{L}\langle], \mathcal{R}\neg$$

$$z \in a, R[z] \subseteq R[y], yS_x a, a \Vdash^{\forall} \neg(B \rhd \bot), y: (B \rhd_y \bot), y: A, y \in R[x], x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots \quad KM1$$

$$yS_x a, a \Vdash^{\forall} \neg(B \rhd \bot), y: (B \rhd_y \bot), y: A, y \in R[x], x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots \quad \mathcal{L}\langle]$$

$$y: \langle]_x \neg(B \rhd \bot), y: (B \rhd_y \bot), y: A, y \in R[x], x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots \qquad \mathbf{y{:}A}, \cdots \Rightarrow \cdots, \mathbf{y{:}A} \quad \mathcal{L}\rhd^{S}$$

$$y: (B \rhd_y \bot), y: A, y \in R[x], x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, \cdots \quad \mathcal{R}\to, \mathcal{R}\neg$$

$$y \in R[x], x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot, y: A \to \neg(B \rhd \bot), \cdots \quad \mathcal{R}\rhd^{S}, \mathcal{L}\neg$$

$$x: A \rhd_x \neg(B \rhd \bot) \Rightarrow x: \neg(A \to \neg(B \rhd \bot)) \rhd_x \bot$$

Figure A4. *Derivation of* KM1.

A.3. $\mathsf{G3ILM_0}$**.** The derivation in $\mathsf{G3ILM_0}$ of the schema $\mathsf{M_0}$ is given in Figure A3. The derivable sequent S occuring in the derivation has shape

$$z \in R[x], z \in R[y], \mathbf{z{:}A}, y \in R[x], \dots \Rightarrow \dots, \mathbf{z{:}A}.$$

A.4. $\mathsf{G3ILKM1}$**.** The derivation in $\mathsf{G3ILKM1}$ of the schema $\mathsf{KM1}$ is given in Figure A4.

A.5. $\mathsf{G3ILW}$**.** For space reasons, we must omit the derivation of the schema W in $\mathsf{G3ILW}$ since it could not be readable in its rendering on screen. In any case, we invite the reader to check by pencil and paper that the principle characterising that extensions is indeed derivable in the calculus we designed.

Acknowledgements. The authors would like to express our sincere gratitude to the two anonymous reviewers for their valuable feedback and insightful comments, which have significantly enhanced the quality of our paper.

The first author wishes to thank Joost Joosten and Luka Mikec for their invaluable support in analysing the literature on interpretability logics, including access to unpublished manuscripts.

Funding. The first author's current position is supported the project SERICS—Security and Rights in CyberSpace PE0000014, financed within PNRR, M4C2 I.1.3, funded by the European Union - NextGenerationEU (MUR Code: 2022CY2J5S); he also acknowledges the EC-COST Action (CA20111) "EuroProofNet". The material discussed here is based on Chapter 5 of his PhD thesis [35], which was supported by a MUR PhD Scholarship (2018-2021).

The second author acknowledges the project "Infinity and Intensionality: Towards A New Synthesis" funded by the Research Council of Norway, and "Modalities in Substructural Logics: Theory, Methods and Applications MOSAIC", funded by the Community Research and Development Information Service (CORDIS) of the European Commission, the MIUR Excellence Department Project awarded to Dipartimento di Matematica, Università di Genova, CUP D33C23001110001 and the "Gruppo Nazionale per le Strutture Algebriche, Geometriche e le loro Applicazioni" (GNSAGA) of the Istituto Nazionale di Alta Matematica (INdAM).

Competing interests. The authors have no relevant financial or nonfinancial interests to disclose.

BIBLIOGRAPHY

[1] Areces, C., Hoogland, E., & de Jongh, D. (2000). Interpolation, definability and fixed points in interpretability logics. In Zakharyaschev, M., Segerberg, K., de Rijke, M., and Wansing, H., editors. *Advances in Modal Logic*, Vol. 2. Stanford, California: CSLI Publications, pp. 53–76.

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

[3] Bilkova, M., Goris, E., & Joosten, J. J. (2004). *Smart Labels*. Liber Amicorum for Dick de Jongh. Amsterdam, NL: Institute for Logic, Language and Computation.

[4] Bilotta, A., Maggesi, M., Perini Brogi, C., & Quartini, L. (2024). Growing HOLMS, a HOL light library for modal systems. In Porello, D., Vinci, C., and Zavatteri, M., editors. *Short Paper Proceedings of the 6th International Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis, OVERLAY 2024, Bolzano, Italy, November 28-29, 2024*, volume 3904 of CEUR Workshop Proceedings, Aachen, DE, pp. 41–48. CEUR-WS.org.

[5] Boolos, G. (1995). *The Logic of Provability*. New York, USA: Cambridge University Press.

[6] Bou, F., & Joosten, J. J. (2011). The closed fragment of IL is PSPACE hard. *Electronic Notes in Theoretical Computer Science*, **278**, 47–54.

[7] de Jongh, D., & Veltman, F. (1990). Provability logics for relative interpretability. In Petkov, P. P., editor. *Mathematical Logic*. Boston, MA: Springer US, pp. 31–42.

[8] ———. (1999). Modal completeness of ILW. In Gerbrandy, D., Marx, M. J., de Rijke, M., and Venema, Y., editors, *Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday*. Amsterdam: Amsterdam University Press.

[9] Girlando, M., Negri, S., & Olivetti, N. (2021). Uniform labelled calculi for preferential conditional logics based on neighbourhood semantics. *Journal of Logic and Computation*, **31**(3), 947–997.

[10] Girlando, M., Negri, S., Olivetti, N., & Risch, V. (2018). Conditional beliefs: from neighbourhood semantics to sequent calculus. *The Review of Symbolic Logic*, **11**(4), 736–779.

[11] Goris, E., Bílková, M., Joosten, J. J., & Mikec, L. (2020). Assuring and critical labels for relations between maximal consistent sets for interpretability logics. arXiv preprint, arXiv:2003.04623.

[12] Goris, E., & Joosten, J. (2008). Modal matters for interpretability logics. *Logic Journal of the IGPL*, **16**(4), 371–412.

[13] Goris, E., & Joosten, J. J. (2020). Two new series of principles in the interpretability logic of all reasonable arithmetical theories. *The Journal of Symbolic Logic*, **85**(1), 1–25.

[14] Hakli, R., & Negri, S. (2012). Does the deduction theorem fail for modal logic? *Synthese*, **187**(3), 849–867.

[15] Hakoniemi, T. A., & Joosten, J. J. (2016). Labelled tableaux for interpretability logics. In *Liber Amicorum Alberti: A Tribute to Albert Visser (Tributes; No. 30)*, Rickmansworth UK: College Publications.

[16] Iwata, S., Kurahashi, T., & Okawa, Y. (2024). The fixed point and the Craig interpolation properties for sublogics of IL. *Archive for Mathematical Logic*, **63**(1), 1–37.

[17] Joosten, J. J., Rovira, J. M., Mikec, L., & Vuković, M. (2024). An overview of Verbrugge semantics, a.k.a. generalised Veltman semantics. In Bezhanishvili, N., Iemhoff, R., and Yang, F., editors. *Dick de Jongh on Intuitionistic and Provability Logics*. Cham: Springer International Publishing, pp. 111–153.

[18] Joosten, J. J., & Visser, A. (2000). The interpretability logic of all reasonable arithmetical theories. *Erkenntnis*, **53**, 3–26.

[19] Kurahashi, T., & Okawa, Y. (2021). Modal completeness of sublogics of the interpretability logic IL. *Mathematical Logic Quarterly*, **67**(2), 164–185.

[20] Maggesi, M., & Perini Brogi, C. (2021). A formal proof of modal completeness for provability logic. In Cohen, L., and Kaliszyk, C., editors. *12th International*

Conference on Interactive Theorem Proving (ITP 2021), volume 193 of Leibniz International Proceedings in Informatics (LIPIcs), Dagstuhl: Schloss Dagstuhl—Leibniz-Zentrum für Informatik, pp. 26:1–26:18.

[21] ———. (2023). Mechanising Gödel-Löb provability logic in HOL light. *Journal of Automated Reasoning*, **67**(3), 29.

[22] Mikec, L. (2022). Complexity of the interpretability logics ILW and ILP. *Logic Journal of the IGPL*, **31**(1), 194–213.

[23] Mikec, L., Joosten, J. J., & Vuković, M. (2020a). A W-flavoured series of interpretability principles. In *Short Papers, Advances in Modal Logic, AiML 2020*, pp. 60–64.

[24] ———. (2020b). On ILWR-frames. Logic and Applications LAP 2020, 50.

[25] Mikec, L., Pakhomov, F., & Vuković, M. (2019). Complexity of the interpretability logic IL. *Logic Journal of the IGPL*, **27**(1), 1–7.

[26] Mikec, L., Perkov, T., & Vuković, M. (2017). Decidability of interpretability logics ILM_0 and ILW^*. *Logic Journal of the IGPL*, **25**(5), 758–772.

[27] Mikec, L., & Vuković, M. (2020). Interpretability logics and generalised Veltman semantics. *The Journal of Symbolic Logic*, **85**(2), 749–772.

[28] Negri, S. (2005). Proof analysis in modal logic. *Journal of Philosophical Logic*, **34**(5), 507–544.

[29] ———. (2014a). Proof analysis beyond geometric theories: From rule systems to systems of rules. *Journal of Logic and Computation*, **26**(2), 513–537.

[30] ———. (2014b). Proofs and countermodels in non-classical logics. *Logica Universalis*, **8**(1), 25–60.

[31] ———. (2017). Proof theory for non-normal modal logics: The neighbourhood formalism and basic results. *IfCoLog Journal of Logics and their Applications*, **4**(4), 1241–1286.

[32] Negri, S., & von Plato, J. (2008). *Structural Proof Theory*. Cambridge, UK: Cambridge University Press.

[33] ———. (2011). *Proof Analysis: A Contribution to Hilbert's Last Problem*. Cambridge, UK: Cambridge University Press.

[34] Okawa, Y. (2024). Unary interpretability logics for sublogics of the interpretability logic IL. *Studia Logica*, **112**(3), 693–721.

[35] Perini Brogi, C. (2022). *Investigations of proof theory and automated reasoning for non-classical logics*. PhD thesis, University of Genoa.

[36] Perini Brogi, C., De Nicola, R., & Inverso, O. (2024). Simpson's proof systems for process verification: A fine-tuning (short paper). In de'Liguoro, U., Palazzo, M., and Roversi, L., editors. *Proceedings of the 25th Italian Conference on Theoretical Computer Science, Torino, Italy, September 11-13, 2024*, volume 3811 of CEUR Workshop Proceedings, Aachen, DE, pp. 292–299. CEUR-WS.org.

[37] Perkov, T., & Vuković, M. (2016). Filtrations of generalized Veltman models. *Mathematical Logic Quarterly*, **62**(4–5), 412–419.

[38] Rovira, J. M., Mikec, L., & Joosten, J. J. (2020). Generalised veltman semantics in agda. In *Short Papers, Advances in Modal Logic, AiML 2020*, pp. 86–90.

[39] Sasaki, K. (2002). A cut-free sequent system for the smallest interpretability logic. *Studia Logica*, **70**(3), 353–372.

[40] Shavrukov, V. Y. (1988). The logic of relative interpretability over Peano arithmetic. Preprint, Steklov Mathematical Institute, Moscow.

[41] ———. (1997). Interpreting reflexive theories in finitely many axioms. *Fundamenta Mathematicae*, **152**(2), 99–116.

[42] Smorynski, C. (1977). *The Incompleteness Theorems*, Studies in Logic and the Foundations of Mathematics, 90, Amsterdam, NL: Elsevier, pp. 821–865.

[43] ———. (2012). *Self-Reference and Modal Logic*. New York, USA: Springer Science & Business Media.

[44] Smullyan, R. M. (1968). Analytic cut. *The Journal of Symbolic Logic*, **33**(4), 560–564.

[45] Solovay, R. M. (1976). Provability interpretations of modal logic. *Israel Journal of mathematics*, **25**(3–4), 287–304.

[46] Švejdar, V. (1991). Some independence results in interpretability logic. *Studia Logica*, **50**(1), 29–38.

[47] Tarski, A. (1933). The concept of truth in formalized languages. In Corcoran, J., editor (1983). *Logic, Semantics, Metamathematics: papers from 1923 to 1938*. United Kingdom: Hackett Publishing Company, pp. 152–278.

[48] Tarski, A., Mostowski, A., & Robinson, R. M. (1953). *Undecidable Theories*, Vol. 13. Amsterdam, NL: Elsevier.

[49] Troelstra, A. S., & Schwichtenberg, H. (2000). *Basic Proof Theory*, 43. Cambridge, UK: Cambridge University Press.

[50] Verbrugge, L. (1992). Verzamelingen-Veltman frames en modellen (set Veltman frames and models). Unpublished manuscript.

[51] Verbrugge, R. L. (2017). Provability logic. In Zalta, E. N., editor. *The Stanford Encyclopedia of Philosophy*. Stanford, California: Metaphysics Research Lab, Stanford University. fall 2017 edition.

[52] Vickers, S. (2014). Continuity and geometric logic. *Journal of Applied Logic*, **12**(1), 14–27.

[53] Visser, A. (1988). Preliminary notes on interpretability logic. Logic group preprint series, 29, Department of Philosophy, University of Utrecht, Utrecht NL.

[54] ———. (1990). Interpretability logic. In Petkov, P.P. editor, *Mathematical Logic*. (pp. 175–209). Boston, MA: Springer US.

[55] ———. (1997). An overview of interpretability logic. Logic Group Preprint Series, 174, Department of Philosophy, University of Utrecht, Utrecht NL.

[56] Vuković, M. (1999). The principles of interpretability. *Notre Dame Journal of Formal Logic*, **40**(2), 227–235.

[57] Zambella, D. (1992). On the proofs of arithmetical completeness for interpretability logic. *Notre Dame Journal of Formal Logic*, **33**(4), 542–551.

IMT SCHOOL FOR ADVANCED STUDIES LUCCA
PIAZZA S.FRANCESCO, 19, 55100 LUCCA LU, ITALY
E-mail: cosimo.perinibrogi@imtlucca.it

UNIVERSITY OF GENOA
VIA BALBI, 5, 16126 GENOVA GE, ITALY
E-mail: sara.negri@unige.it

AIX-MARSEILLE UNIVERSITY
JARDIN DU PHARO, 58 BOULEVARD CHARLES LIVON
13007 MARSEILLE, FRANCE
E-mail: nicola.olivetti@univ-amu.fr