Growing HOLMS: A Verified Automated Prover for Grzegorczyk Logic in HOL Light

Antonella Bilotta[1(✉)], Marco Maggesi[2], and Cosimo Perini Brogi[3]

[1] Scuola Normale Superiore di Pisa, Pisa, Italy
antonella.bilotta@sns.it
[2] University of Florence, Florence, Italy
marco.maggesi@unifi.it
[3] IMT School for Advanced Studies Lucca, Lucca, Italy
cosimo.perinibrogi@imtlucca.it

Abstract. This paper presents a certified theorem prover for Grzegorczyk logic (Grz) implemented in the general-purpose proof assistant HOL Light. Our prover builds on original HOL Light formalisations of modal adequacy for Grz with respect to finite partially ordered frames, and on the standard full and faithful translation of Grz into Gödel–Löb logic (GL). This formalised embedding allows us to extend the range of modal systems supported by the HOLMS library for automated modal reasoning, and constitutes a new methodology experimented in our framework, being the first logic added to the library through a modal translation.

The deductive engine performs an automated proof search in the labelled sequent calculus for GL. When the proof search on the translated formula succeeds, the system returns a HOL Light theorem certifying provability of the original Grz formula. When proof search terminates negatively, the system constructs a verified GL countermodel and thus certifies that the original formula is not provable in Grz.

Keywords: Automated Theorem Proving · Modal Reasoning · HOL Light · Grzegorczyk Logic · Gödel-Löb Logic · Modal Embedding · Logical Verification

1 Introduction

Recent formalisations of modal logic have progressed from verifying isolated systems to building comprehensive, modular tools.[1] The HOL Light Library for Modal Systems (HOLMS) [13,14,17] exemplifies this trend, providing a unified framework for reasoning about and within seven normal modal logics (K, T, B, K4, S4, S5, and GL) using the HOL Light proof assistant [54,55].

HOLMS employs a dual methodology: a *deep embedding* of axiomatic calculi and relational semantics to prove modal adequacy theorems, alongside a *shallow*

[1] See, e.g., [38–40,43,61].

A. Biere et al. (Eds.): IJCAR 2026, LNAI 16688, pp. 414–435, 2026.
https://doi.org/10.1007/978-3-032-32589-1_25

embedding of labelled sequent calculi to facilitate automated reasoning. Consequently, the framework yields both verified theorems within the modal systems and certified countermodels upon proof search failure.

Our Contribution. We integrate HOLMS with a first significant example of certified modal translations. The intent is to extend the range of modal logic implemented in the library and experiment with applications of modal embeddings to modular automated reasoning in HOL Light. Here, we instantiate the methodology on Grzegorczyk logic (Grz) [46], a system bridging classical modal logic, intuitionistic mathematics, and topology [26].

Automated reasoning for Grz is traditionally hindered by its characteristic frame condition – specifically, the *weakly Noetherian* property (converse weak well-foundedness). Consequently, to implement a sequent calculus for Grz constructively typically demands non-trivial, ad hoc design choices to handle this condition, such as in [31] or in [73].

Rather than complicating the uniform architecture of HOLMS with such mechanisms, we circumvent the weak Noetherian condition entirely by formalising the Kuznetsov-Goldblatt-Boolos theorem [20,21]. In particular, we mechanise a full and faithful embedding of Grz into Gödel-Löb provability logic (GL) in two steps:

- implementing the standard *splitting translation* within HOL Light;
- reducing validity problems in Grz to corresponding validity problems in GL.

By doing so, we directly reuse HOLMS's existing verified decision procedure for GL [59] to decide Grz-theoremhood automatically.

More Technically. Consistent with HOLMS's architectural principles, we first provide a deep embedding of Grzegorczyk axiomatisation ($\mathbb{GRZ}$) as the extension of the minimal normal system $\mathbb{K}$ by the Grzegorczyk schema:

$$\Box(\Box(A \to \Box A) \to A) \to A.$$

We then prove that this axiomatisation is equivalent to $\mathbb{S}4\mathbb{GRZ}$ obtained by adding the same schema to the system $\mathbb{S}4$.

On the semantic side, we formalise the classes of RTWN (Reflexive, Transitive, Weakly Noetherian) frames and RATF (Reflexive, Antisymmetric, Transitive, Finite) frames. By proving the soundness and completeness of the axiomatic systems with respect to these classes, we integrate Grz into the library's existing semantic hierarchy alongside S4 and GL.

Whereas HOLMS typically relies on a shallow embedding of labelled sequents to drive automation, we adopt a translational approach for Grz to bypass the proof-theoretic complexities of its calculi. We implement the standard *splitting translation*, denoted `TRANSL` or $(\cdot)^+$, as a recursive function mapping modal formulas to modal formulas:

$$(\Box A)^+ := A^+ \wedge \Box(A^+)$$

This translation "internalises" reflexivity into the irreflexive framework of GL. We formally prove fullness and faithfulness of the embedding (`GRZ_TRANSL`), establishing that $\mathbb{GRZ} \vdash A$ if and only if $\mathbb{GL} \vdash A^+$. This result allows us to repurpose the original, certified decision procedure for GL (`HOLMS_RULE`), to decide validity in Grz.

A distinct advantage of the HOLMS automated reasoning is its capacity to produce certified countermodels for unprovable formulas in normal systems present in the codebase, including GL. Our approach preserves this constructive feature via a certified model transformation, executed in three steps:

1. When GL decision procedure refutes a translated formula A^+, the original tactic `HOLMS_BUILD_COUNTERMODEL` returns a finite, irreflexive, transitive countermodel for A^+;
2. We formally define a transformation, `R_PLUS`, which computes the reflexive closure of such models.
3. We prove that this transformation maps finite GL-models for A^+ to finite Grz-models for A within the class RATF.

Thus, a failed proof search yields a concrete, verified Grz-countermodel for the original formula.

This automation of Grz-reasoning paves the way for extending the library's certification guarantees to intuitionistic validity via the Gödel-McKinsey-Tarski translation [41,60], which is an additional instance of the method of verified modal embeddings we advocate for.

Structure of the Paper. The remainder of this paper is organised as follows. Section 2 provides background on the HOLMS library, detailing its existing deep and shallow embeddings. Section 3 presents our formalisation of Grzegorczyk logic, including the axiomatic definitions and the proof of modal adequacy with respect to the standard frame classes. Section 4 details the formal verification of the Kuznetsov-Goldblatt-Boolos theorem, establishing the faithfulness of the splitting translation. In Sect. 5, we describe the decision procedure and the countermodel transformation R^+, by also presenting concrete examples of application; furthermore, we discuss aspects of the computational costs involved in the implementation of Grz-reasoning via modal embedding in GL. Section 6 concludes the main body of the paper with a summary and directions for future work, particularly regarding other modal companions.

A preprinted version of this work [15] includes two supplementary appendices to which we refer in the following. These detail the decision algorithm used in HOLMS, and the labelled sequent calculi underlying the procedure, respectively.

Source Code. The HOLMS codebase is freely available on GitHub, with the version presented here permanently archived as release v2026-05-24. Additionally, the HOLMS project webpage [16] offers links to practical and theoretical documentation, concerning HOL Light and the framework.

The existing core library was extensively detailed in Appendix A of our previous work [14], while the updated code extends this repository with three additional Ocaml files:

- `grz.ml` contains the formalisation of Grz and the implementation of its decision procedure via modal translation described in Sects. 3–5;
- `translations.ml` defines and exemplifies the splitting translation, as described in Sect. 4;
- `grz_tests.ml` provides some tests of the decision procedure for Grz, alongside those discussed in Sect. 5.

In the following pages, hyperlinks – denoted by the symbol – point to code snippets relevant for the presentation.

2 Background on the Core Library

The HOLMS library provides a modular environment for formal reasoning about seven modal logics (K, T, K4, S4, B, S5, GL) within the HOL Light proof assistant. Its architecture is characterised by three interconnected presentations of modal logics: axiomatic calculi, relational semantics, and labelled sequents.

The library's principal innovation lies in its modularity, structured around Strachey's distinction between *parametric polymorphism* and *ad hoc polymorphism* [81]. This methodological choice clearly separates the verified infrastructure common to all systems from the bespoke components required for individual logics during proof development.

Parametric polymorphism denotes constructions where a single, uniform implementation serves all logic instances. In HOLMS, this encompasses the modal syntax, the deducibility relation, and the foundational lemmas underlying completeness proofs, such as the standard model construction and the truth lemma.

Conversely, *ad hoc polymorphism* describes bespoke implementations tailored to specific logics, with the appropriate variant selected at the meta-theoretical level. This category includes logic-specific accessibility relations (e.g., the irreflexive and transitive relation for GL), individual decision procedures, and the correspondence lemmas for distinct frame properties.

This section discusses the key aspects of the library; for a more detailed account, we refer the reader to [14].

2.1 Deep Embedding: Axiomatic and Semantic Presentations

The library employs a deep embedding for both the syntax and relational semantics of normal modal logics.

The modal language $\mathcal{L}$ is defined via an inductive type `:form`. Deducibility is captured by a ternary relation $\mathcal{S}.\mathcal{H} \vdash A$, parameterised by a set of axiom schemas $\mathcal{S}$, which modularly extends the standard calculus $\mathbb{K}$ [33] into logics such as T, K4, S4, B, S5, and GL according to standard literature [18,19].[2]

[2] The notation `[S.{}|~ a]` denotes thus a term of HOL Light metalanguage representing a theorem of the object modal logic axiomatised by the schemas in $\mathcal{S}$.

Semantic validity is defined over Kripke frames $\langle W, R \rangle$ via a standard forcing relation $\Vdash$ (the formal predicate `holds` ↗) induced by an evaluation function $v : \texttt{Atoms}_{\mathcal{L}} \times W \longrightarrow \{0, 1\}$ [4]. Correspondence theory then maps these frames to specific logics.

While the notions of frames and characteristic classes are parametric, specific correspondence lemmas – such as `MODAL_TRANSNT` ↗ for GL – are ad hoc implementations tailored to the distinct relational properties of each system.

In the codebase, `CHAR S` ↗ denotes the class of frames satisfying the characteristic property of $\mathcal{S}$, and `APPR S` ↗ denotes the class of finite frames in which any theorem of the axiomatic calculus defined by $\mathcal{S}$ is valid. The lemma `APPR_EQ_CHAR_FINITE` ↗ proves that the class of appropriate frames `APPR S` coincides with the characteristic class `CHAR S` restricted to finite frames.

The modular formalisation of modal adequacy (soundness and completeness) connects the seven axiomatic systems with their semantic presentations. Completeness is established via a principled modularisation of Boolos's strategy from [21], first mechanised in [58], decomposing the proof into:

- *Parametric components*, providing a uniform implementation applicable to all systems. These encompass construction of a `GEN_STANDARD_MODEL` ↗ from maximal consistent lists of formulas, proof of a `GEN_TRUTH_LEMMA` ↗, and identification of a generic counterworld via `GEN_COUNTERMODEL_ALT` ↗.
- *Ad hoc components*, necessitating logic-specific customisation. These include the definition of standard accessibility relations, as `GL_STANDARD_REL` ↗, the `ACCESSIBILITY_LEMMA` instances (e.g., `GL_ACCESSIBILITY_LEMMA` ↗), and theorems that verify that the constructed countermodel inhabits the correct characteristic frame class (e.g., `ITF_MAXIMAL_CONSISTENT` ↗) .

This dual approach aims to guarantee the meta-theory remains scalable: extending the library to a new logic primarily requires supplying the requisite ad hoc relational definitions while reusing the existing parametric infrastructure.

2.2 Shallow Embedding: Labelled Sequents and Automation

For automated reasoning across its seven supported systems, HOLMS employs a shallow embedding of G3-style labelled sequent calculi [77,85], optimised according to the methodology of [66–68]. Definitions of these calculi are collected in Appendix B of [15].

This embedding maps labelled sequents directly onto the HOL Light goal stack: formulas on the left-hand side of a sequent manifest as native hypotheses, while those on the right-hand side are encoded as subgoals in disjunctive form. In this architecture, a generic labelled sequent $\mathcal{R}, \Gamma \Rightarrow \Delta$ is encoded such that:

- The relational atoms in $\mathcal{R}$ and the forcing atoms in Γ serve as HOL Light assumptions;
- The succedent Δ is encoded as the disjunction of its constituent forcing atoms within the current HOL Light goal.

This shallow embedding strategy permits the direct reuse of the meta-logic's native deductive engine. Notably, classical propositional connectives require no bespoke encoding: they are handled directly by standard HOL Light tactics such as `IMP_TAC`, `CONJ_TAC`, and `DISJ_TAC`.

Modal rules, by contrast, are reconstructed at the meta-level by appealing to the forcing relation on relational frames.[3] Within this framework, each modal rule is implemented as a HOL Light tactic, even those capturing the semantic properties summarised in Table 3 of Appendix B from [15]. For instance, `BOX_RIGHT_THEN` tactic realises the $\Box$-rule in the consequent.[4]

The labelled sequent calculus we employ for reasoning about Grzegorczyk logic, via modal embedding to GL, is G3KGL. Its rules are summarised in Table 4 of Appendix B from [15], and we shallowly embedded them in HOL Light in previous work [59]. The automated decision procedure is then realised by `HOLMS_RULE` and the associated tactic `HOLMS_TAC`. Such a realisation implements the the goal transitions shown in Fig. 1 of Appendix A from [15], which we rephrase as follows:

- **Starting Goal**: The user supplies a modal formula A, a set of hypotheses $\mathcal{G}$, and specifies one of seven modal logics $* \in \{\text{K, T, K4, S4, B, S5, GL}\}$. Each logic is characterised by a specific set $\mathcal{S}$ of axiom schemas.
- **Goal A:** The provability question is first translated into a semantic one: *is A valid within the class of Kripke frames characteristic of $\mathcal{S}$?* Our mechanised adequacy theorems justify this translation.
- **Goal B:** Using the shallow embedding and its internalisation of Kripke semantics, the problem is then recast as checking whether *A can be formally derived in the labelled calculus associated with $*$.*
- **Automated Proof Search:** We implement a root-first proof search for each calculus, which yields a decision procedure for $*$ [67]. The outcome branches as follows:
  - **Success branch:** When the search reaches initial sequents, `HOLMS_TAC` generates a shallowly embedded derivation of A which HOL Light certifies as a fully trusted proof of validity.
  - **Failure branch:** If the search fails, `HOLMS_BUILD_COUNTERMODEL` extracts a candidate countermodel from the HOL Light goal-stack.[5] Finally, `CERTIFY_COUNTERMODEL_TAC` calls HOL Light to verify that the candidate falsifies the input formula A at its root.

Currently, HOLMS provides decision procedures for four logics (K, T, B, and GL) – each halts by either proving validity or producing a certified finite countermodel – and semi-decision procedures for K4, S4, and S5.[6] In Appendix

[3] Refer to [68] for a detailed discussion of the method.

[4] Crucially, weakening, contraction, and cut rules are not introduced axiomatically but emerge naturally from HOL Light's underlying logical engine.

[5] This extraction mirrors the countermodel construction from failed proof searches well known in proof theory [67,72,82].

[6] For K4 and its extensions, the algorithm may terminate owing to a bound on the size of the generated countermodel, which prevents infinite loops caused by the

A of [15] we give a more detailed account of the design and functioning of the certified theorem-proving mechanism within HOLMS.

3 Grzegorczyk Logic, Formalised

We present the deep embedding of Grzegorczyk logic (Grz) in HOLMS, covering both its axiomatic derivability relation and its characteristic frame classes, following the methods of Sect. 2.1. The syntactic and semantic presentations are then bridged by formalising soundness and completeness for Grz, following the strategy developed in [21, § 12]. Observe that this is a different strategy from the modular one developed in [21, § 5] and implemented in the previous version of the HOLMS core library.

3.1 Axiomatic Systems for Grz

We introduce two distinct axiomatisations of Grz and formally prove their deductive equivalence. The core Grzegorczyk axiom schema, standardly denoted Grz, is

$$\Box(\Box(A \to \Box A) \to A) \to A.$$

In our code this is captured by the definition `GRZ_AX`. The purely axiomatic calculus $\mathbb{GRZ}$ is then obtained by instantiating the general derivability predicate $\mathcal{S}.\mathcal{H} \vdash A$ with $\mathcal{S} := \{\mathsf{Grz}\}$, side-stepping at once the debate about the failure of necessitation [50].

Definition 1. *The predicate $\mathsf{Grz}.\mathcal{H} \vdash A$, denoting derivability of a formula A from a set of hypotheses $\mathcal{H}$ in $\mathbb{GRZ}$, is inductively defined as follows:*

- *for each instance A of a schema of classical propositional logic, $\mathsf{Grz}.\mathcal{H} \vdash A$;*
- *for each instance A of schema $\mathsf{K} := \Box(B \to C) \to \Box B \to \Box C$, $\mathsf{Grz}.\mathcal{H} \vdash A$;*
- *for each instance A of schema Grz, $\mathsf{Grz}.\mathcal{H} \vdash A$;*
- *for each $A \in \mathcal{H}$, $\mathsf{Grz}.\mathcal{H} \vdash A$;*
- *if $\mathsf{Grz}.\mathcal{H} \vdash B \to A$ and $\mathsf{Grz}.\mathcal{H} \vdash B$, then $\mathsf{Grz}.\mathcal{H} \vdash A$;*
- *if $\mathsf{Grz}.\varnothing \vdash A$, then $\mathsf{Grz}.\mathcal{H} \vdash \Box A$ for any set of formulas $\mathcal{H}$.*

Alternatively, Grz can be presented as an extension of the modal logic $\mathbb{S}4$. We define the calculus $\mathbb{S}4\mathbb{GRZ}$ (`S4GRZ_AX`) analogously to Definition 1, but with the axiom schemas supplemented by $\mathsf{4} := \Box A \to \Box\Box A$ (`FOUR_SCHEMA`) and $\mathsf{T} := \Box A \to A$ (`T_SCHEMA`); thus the axiomatic base consists of classical logic, K, T, $\mathsf{4}$, and Grz. Formalising the known equivalence between the two axiomatisations (`GRZ_EQ_S4GRZ`) allows us to work interchangeably with whichever presentation is most convenient. Accordingly, we prove the adequacy theorem for $\mathbb{S}4\mathbb{GRZ}$.

interaction between transitive closure and logical rules. However, the present tactic does not yet certify this bounded countermodel; established termination results [37] are planned for integration into future versions of the library.

3.2 Relational Frames for Grz

We now turn to the semantic characterisation of Grzegorczyk logic. A Kripke frame $\langle W, R\rangle$ validates Grz precisely when R is reflexive, transitive, and *weakly Noetherian*: the converse R^{-1} is well-founded in the sense that every non-empty subset of W contains an R-maximal element. To formalise this we introduce the definition of weakly well-founded relations `WWF` ↗. For a relation $\prec$,

$$\texttt{WWF}(\prec) \;:=\; \forall P \subseteq \mathrm{fld}(\prec).\; P \neq \emptyset \implies \exists x \in P.\; \forall y \in P.\; \neg(y \prec x \wedge y \neq x).$$

states that every non-empty subset of the field of $\prec$ contains a minimal element w.r.t. the strict version of $\prec$.[7] Equivalent formulations that facilitate different styles of reasoning are provided by `WWF_EQ` ↗, `WWF_IND` ↗, and `WWF_ALT` ↗.

A relation that is reflexive, transitive, and antisymmetric is a partial order; we capture this with `POSET_EQ` ↗. The key fact that any finite, non-empty partial order is weakly well-founded is established in `POSET_WWF` ↗, together with its set-theoretic counterpart `POSET_WWF_ALT` ↗.

With these foundations we define the specific frame classes relevant to Grz.

Definition 2 (`RTWN_DEF` ↗)**.** *A frame $\langle W, R\rangle$ belongs to the class RTWN if R is reflexive, transitive, and the converse R^{-1} is weakly well-founded (`WWF`(R^{-1})).*

Definition 3 (`RATF_DEF` ↗)**.** *A frame $\langle W, R\rangle$ belongs to the class RATF if R is reflexive, antisymmetric, transitive, and W is finite.*

By `POSET_WWF_ALT` ↗, each finite partial order satisfies the weakly Noetherian condition, so we can prove RATF ⊆ RTWN (`RATF_SUBSET_RTWN` ↗).

3.3 Formal Proof of Adequacy

Following Boolos [21, § 12], we do not apply the general frame strategy of Sect. 2.1 to obtain adequacy for $\mathbb{GRZ}$. Instead, using the logical equivalence of $\mathbb{GRZ}$ and $\mathbb{S}4\mathbb{GRZ}$, we proceed in three steps:

1. RATF ⊆ RTWN (`RATF_SUBSET_RTWN` ↗)
2. RTWN is sound for $\mathbb{S}4\mathbb{GRZ}$ (Lemma 1);
3. RATF is complete for $\mathbb{S}4\mathbb{GRZ}$ (Theorem 1).

Soundness. The core of the soundness argument is `MODAL_RTWN` ↗, which proves that the schema Grz is valid in every frame whose accessibility relation is reflexive, transitive, and weakly Noetherian.[8] Combined with the correspondence lemmas of T (reflexivity, `MODAL_REFL` ↗) and 4 (transitivity, `MODAL_TRANS` ↗), we obtain that the class RTWN validates all axioms of $\mathbb{S}4\mathbb{GRZ}$ (`RTWN_SUBSET_CHAR_S4GRZ` ↗). Instantiating the generic soundness theorem of the library then yields:

[7] Notice that $\mathrm{fld}(\prec) = W$ whenever the relation $\prec$ is reflexive; thus we can define `WWF`$(\prec)$ on subsets of W.

[8] The proof relies on the induction principle `WWF_IND` to exclude infinite, strictly ascending R-chains.

Lemma 1 (`S4GRZ_RTWN_VALID` ↗). *If* $\mathsf{S4GRZ}.\mathcal{H} \vdash A$ *and every formula in* $\mathcal{H}$ *is valid on RTWN, then RTWN* $\vDash A$.

From this, soundness for RATF follows immediately (`S4GRZ_RATF_VALID` ↗). As a consistency check, `S4GRZ_CONSISTENT` ↗ is derived trivially.

Completeness. We prove completeness w.r.t. the finite frames in RATF.

Theorem 1 (`S4GRZ_COMPLETENESS_THM` ↗). *If* RATF $\vDash A$, *then* $\mathsf{S4GRZ}.\varnothing \vdash A$.

Proof (Sketch). By contraposition, suppose A is not provable in $\mathbb{S}4\mathbb{GRZ}$. The worlds of the countermodel we are aiming for are maximal consistent lists of formulas built from the subformulas of A together with their $\Box$-subformulas (`S4GRZ_STANDARD_WORLDS_DEF` ↗).

A Lindenbaum construction (`GRZ_NONEMPTY_MAXIMAL_CONSISTENT` ↗) produces such a list l containing $\neg A$. The accessibility relation `S4GRZ_STANDARD_REL_DEF` ↗ is obtained from a base relation `Q_REL` ↗ by enforcing antisymmetry ($l\ Q\ l' \wedge l'\ Q\ l \implies l = l'$); `S4GRZ_STANDARD_REL_PROP` ↗ proves it reflexive, antisymmetric, and transitive. Since the list of subformulas is finite, the frame is finite, hence belongs to RATF. The truth lemma (`S4GRZ_TRUTH_LEMMA` ↗) then yields $l \nVdash A$ as required.[9]

4 Certified Splitting Translation

The core formalised result in this section is that every theorem of $\mathbb{GRZ}$ can be fully and faithfully translated into a theorem of $\mathbb{GL}$. Since Grzegorczyk logic is complete for finite posets (RATF) and Gödel-Löb logic is complete for irreflexive, transitive and finite frames (ITF), the proof proceeds by formally converting the accessibility relations back and forth and implementing the standard splitting translation $(\cdot)^+$ of $\mathbb{GRZ}$-theorems into $\mathbb{GL}$-theorems.

4.1 Translation Function

Let us recall first the translation between formulas.

Definition 4 (Splitting translation, `TRANSL` ↗**).** *The function* $(\cdot)^+ : \mathit{Form}_{\mathcal{L}} \to \mathit{Form}_{\mathcal{L}}$ *is defined by induction on the size of a formula:*

$$
\begin{aligned}
(\bot)^+ &= \bot,\\
(p)^+ &= p \text{ for any atom } p,\\
(\neg A)^+ &= \neg(A^+),\\
(A \circ B)^+ &= A^+ \circ B^+ \text{ for any binary operator } \circ,\\
(\Box A)^+ &= \Box(A^+) \wedge A^+.
\end{aligned}
$$

[9] By `POSET_WWF_ALT`, the countermodel we get from the Lindenbaum construction is indeed in RTWN, so that Boolos' proof could be even simplified, as noticed by a reviewer.

Notice that the effective, non-trivial translation happens for the modal clause: $(\Box A)^+$ coincides with $\boxdot(A^+)$, where $\boxdot B := \Box B \wedge B$ is the usual "box-and" modality.

This translation precisely bridges the semantic gap between the two logics. In $\mathbb{GRZ}$ (and $\mathbb{S}4\mathbb{GRZ}$), the $\Box$ is reflexive: what is necessary is also actually true. In $\mathbb{GL}$, by contrast, $\Box$ represents formal provability and is strictly irreflexive – because of Gödel's incompleteness theorems [78,79]. Simulating a $\mathbb{GRZ}$ box-modality inside $\mathbb{GL}$ therefore requires explicitly asserting the actual truth of the formula together with its modal truth. The wider significance of this translation for the study of formal provability via Grz is discussed in [20,21].

4.2 Relational Transformers

To mediate between the frame classes ITF and RATF we define two operators that geometrically manipulate the accessibility relation R of a frame.

- R^+ is the reflexive closure operator (`R_PLUS_DEF`): $R^+ := R \cup \{\langle x, x\rangle : x \in W\}$.
- R^- is the irreflexive core operator (`R_MIN_DEF`): $R^- := R \setminus \{\langle x, x\rangle : x \in W\}$.

Basic sanity checks are immediate: if R is already irreflexive, R^- leaves it unchanged (`IRREFL_THEN_EQ`); conversely, a reflexive relation can be exactly recovered by taking its irreflexive core and then adding back all reflexive arrows (`REFL_THEN_EQ`).

4.3 Kuznetsov-Goldblatt-Boolos Theorem

We now establish the mathematical heart of the translation: the syntactic mapping $(\cdot)^+$ mirrors the geometric lifting $R \mapsto R^+$ in the sense that evaluating a translated formula A^+ on an ITF frame yields the same truth value as evaluating the original formula A on the reflexively closed version of that very frame.

Lemma 2 (`GRZ_TRANSL_LEMMA`, `ITF_IMP_R_PLUS_RATF`). *Let $\langle W, R, v\rangle$ be a model with $\langle W, R\rangle \in$ ITF. For any world $w \in W$ and formula $A \in \mathit{Form}_{\mathcal{L}}$:*

i) $w \Vdash_{\langle W,R,v\rangle} A^+ \iff w \Vdash_{\langle W,R^+,v\rangle} A$;
ii) $\langle W, R^+, v\rangle \in$ RATF.

The next step links validity on RATF frames to provability in $\mathbb{GL}$: we prove that a formula A is valid on all finite reflexive posets if and only if A^+ is a theorem of $\mathbb{GL}$ (`RATF_TRANSL`).[10] With this, we obtain the complete formalisation of the theorem by Kuznetsov, Goldblatt, and Boolos.

[10] The proof uses a two-way contrapositive argument, combining Lemma 2 with the existing soundness and completeness theorems for $\mathbb{GL}$ (`GL_ITF_VALID` and `GL_COMPLETENESS_THM`).

Theorem 2 (Kuznetsov-Goldblatt-Boolos, `GRZ_TRANSL`). *For every formula* $A \in \mathit{Form}_{\mathcal{L}}$,

$$\mathsf{Grz}.\varnothing \vdash A \iff \mathsf{GL}.\varnothing \vdash A^{+}.$$

Proof (Sketch). Combine `RATF_TRANSL` with the adequacy theorem for $\mathbb{GRZ}$ (`GRZ_ADEQ_THM`) to eliminate the semantic middleman from Lemma 2.

5 Grzegorczyk Logic, Automated

The theoretical results established in the preceding sections directly yield an automated decision procedure for Grzegorczyk logic. By leveraging the full and faithful embedding of Theorem 2, we reuse the existing verified solver for Gödel-Löb logic within HOLMS.

The pipeline underlying our novel HOL Light tactic, `GZR_TAC`, integrates into the general `HOLMS_TAC` as follows, with each stage illustrated by examples:

1. We intercept a goal of the form $\mathsf{Grz}.\varnothing \vdash A$ and rewrite it into $\mathsf{GL}.\varnothing \vdash A^{+}$, in virtue of the formal proof of the core embedding theorem `GRZ_TRANSL`;

   Example 1:

   Suppose that we wish to check whether Grz proves:
   $(\star)$ $\quad \mathit{contingent}(a) \to \Diamond(\mathit{penultimate}(a) \vee \mathit{penultimate}(\neg a))$[a]
   where $\mathit{contingent}(a) := \Diamond a \wedge \Diamond(\neg a)$ and
   $\mathit{penultimate}(a) := a \wedge \Diamond(\neg a) \wedge \Box(\neg a \to \Box\neg a)$, so the formula is:
   $(\star)\ \Diamond a \wedge \Diamond\neg a \to \Diamond(a \wedge \Diamond\neg a \wedge \Box(\neg a \to \Box\neg a) \vee \neg a \wedge \Diamond\neg\neg a \wedge \Box(\neg\neg a \to \Box\neg\neg a))$

   [a] This corresponds to the axiom $\mathsf{Grz}^{\star}$ from [1].

   Example 2:

   Suppose that we wish to check if Grz proves the convergence axiom:

   $$(.2)\ \Diamond\Box a \to \Box\Diamond a$$

2. We traverse the abstract syntax tree of the formula, replacing every $\Box B$ with $\boxdot B$, mathematically computing the shape of the splitting translation of A;

   Example 1 (Continued):

   We convert $(\star)$ into its GL-equivalent, which we denote by $(\star)^{+}$:

   $$\begin{aligned} &\neg\boxdot\neg a \wedge \neg\boxdot\neg\neg a \\ &\quad\to\ \neg\boxdot\neg\Big(a \wedge \neg\boxdot\neg\neg a \wedge \Big(\boxdot\big(\neg a \to \boxdot\neg a\big) \\ &\qquad\qquad \vee\ \big(\neg a \wedge \neg\boxdot\neg\neg\neg a \wedge \boxdot(\neg\neg a \to \boxdot\neg\neg a)\big)\Big)\Big) \end{aligned}$$

Example 2 (Continued):

We convert $(.2)$ into its $\mathbb{GL}$-equivalent, which we denote by $(.2)^+$:

$$\neg \boxdot (\neg \boxdot (a)) \rightarrow \boxdot(\neg \boxdot (\neg a)).$$

3. We invoke `GL_TAC` 🔗, delegating proof search to the existing decision procedure for the shallow embedding of G3KGL;

Example 1 (Continued)

The proof search operates on the computed conjunctive normal form of $(\star)^+$ For brevity, we provide the following statistics about the resulting formula:

Metric	Formula Size	Number of Clauses	Clause Size (Max)	Clause Size (Min)
Value	4359	44	7	3

Example 2 (Continued)

The proof search operates on the computed conjunctive normal form of $(.2)^+$. For brevity, we provide the following statistics about the resulting formula:

Metric	Formula Size	Number of Clauses	Clause Size (Max)	Clause Size (Min)
Value	59	4	2	2

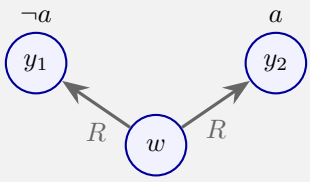
4. If proof search succeeds, the system returns a HOL Light theorem certifying the validity of A^+. Upon failure, `HOLMS_BUILD_COUNTERMODEL` 🔗 generates a certified countermodel $\langle W, R, v\rangle$ for A^+ in $\mathbb{GL}$.

Example 1 (Continued)

The tactic succeeds for $(\star)^+$, completing the proof-search in 1.58 seconds, on mid-level personal computer.

Example 2 (Continued)

The tactic fails for $(.2)^+$, and the system returns the countermodel built at the end of the proof search, graphically rendered here as:



5. `GRZ_CONV` certifies the validity of the whole decision procedure by exploiting Kuznetsov-Goldblatt-Boolos Theorem (`GRZ_TRANSL`).

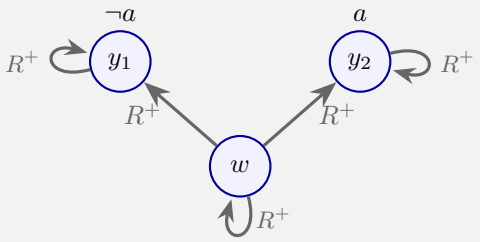
Example 1 (Completed)

The system returns a new HOL Light theorem stating that the original formula (1) is provable in Grz, that is:

```
|- [GRZ_AX . {}
    |~ Contingent(Atom "a") -->
          Diamond (Penultimate(Atom "a") \/
                   Penultimate(Not Atom "a"))]
```

Example 2 (Completed)

The $\mathbb{GL}$-countermodel is transformed into a Grz-countermodel via the transformer $(\cdot)^{-}$, and the schema (.2) is automatically model-checked against the resulting structure:



which is formally proven to be a countermodel by `GRZ_CONV` to $\Diamond\Box a \to \Box\Diamond a$.

5.1 Scalability, Complexity, and Practical Boundaries

The viability of leveraging translation-based frameworks in modal proof theory depends heavily on the trade-offs between conceptual modularity and computational efficiency.

The primary merit of our approach lies in its architectural extensibility: the splitting translation enables the re-use of the existing automated solver in HOLMS. This modularity yields a significant proof-theoretic dividend: it provides a uniform pipeline to reason about any intermediate logic or translation that embeds into Grz.

A prime exemplar of this broader applicability is the Gödel-McKinsey-Tarski translation, which embeds intuitionistic propositional logic (IPC) into Grz.[11]

[11] We implemented the translation in the file `god_transl.ml` of our library, without discussing its computational ramifications here.

Thus HOLMS inherits a native and verified pipeline from IPC to GL, suggesting that our method is not an isolated curiosity, but a theoretically scalable paradigm for executing cross-logic proof searches within a singular target back-end.

Despite these structural advantages, the naïve application of the splitting translation introduces a severe structural bottleneck regarding formula size. Because the translation duplicates subformulas, it suffers from severe spatial inefficiency. This subformula duplication induces exponential growth in the total size of the formula relative to its modal depth. Consider a worst-case syntactic benchmark consisting of n nested modalities, $B = \Box^n A$. Under the splitting map, the size of the translated formula behaves according to the recurrence relation $\text{size}((\Box^k A)^+) = \text{size}((\Box^{k-1} A)^+) + \text{size}(\Box(\Box^{k-1} A)^+) + 1$ yielding an asymptotic complexity strictly bounded by $\mathcal{O}(2^n)$. For deeply nested formulas, this exponential blow-up rapidly exhausts memory bandwidth and generates massive inputs for the underlying back-end.

The validity problem for Grz is well-established in the literature as PSPACE-complete; because GL shares this identical complexity class [84], the exponential blow-up induced by the translation artificially exacerbates the computational cost for the target GL prover.[12] This significantly impacts the scalability of the method during intense proof searches, such as those arising from the modal analysis of intuitionistic statements. Following the insightful remarks of a reviewer, in future work, we intend to explore the possibility of restoring the practical advantages of the shared PSPACE complexity class.

6 Conclusions

We have presented a methodology for exploring modal translations within the HOL Light Library for Modal Systems (HOLMS). We illustrated this approach through the splitting translation of Grzegorczyk logic (Grz), a central system in mathematical modal logic [26]. This demonstrates that a modular formalisation strategy effectively circumvents the proof-theoretic intricacies arising from non-trivial frame conditions, notably weak Noetherianity.

Our development comprises three components. First, we provided a deep embedding of the axiomatic systems for Grz, establishing soundness and completeness w.r.t. RTWN (reflexive, transitive, weakly Noetherian) and RATF (reflexive, antisymmetric, transitive, finite) frames. Second, we formalised the Kuznetsov–Goldblatt–Boolos theorem, reducing $\mathbb{GRZ}$-theoremhood checking to the existing, verified decision procedure for $\mathbb{GL}$. Finally, we mechanised the model transformation underlying the semantic proof of faithfulness, so that failed proof attempts yield certified RATF countermodels derived via the reflexive closure of the GL-countermodel generated by HOLMS tactic for Gödel-Löb logic.

[12] GL proof search can theoretically be executed using polynomial space relative to its input; solving a translated formula B should thus remain tractable, but the translation itself shifts the problem from $\text{poly}(|B|)$ to $\text{poly}(2^{|B|})$, effectively degrading the practical performance to EXPSPACE boundaries.

This extension consolidates HOLMS as a unified platform for modal metatheory and automated reasoning, establishing a foundation for further developments along different directions.

Future Work. The success of this first experiment with orthogonal extensions of HOLMS via modal embeddings invites us to plan optimisations (inspired by Tseitin transformations [83]) of the decision procedures in terms of performance and usability of the framework as the library grows. A promising complementary direction is the formalisation of a shallow embedding of Dyckhoff and Negri's non-harmonic labelled calculus G3Grz [31], which would provide a jump from exponential complexity to polynomial constraint on the proof search, and would also demonstrate HOLMS's capabilities in unifying disparate reasoning paradigms. We also aim to certify the Gödel–McKinsey–Tarski translation, reflecting the established connections between Grz and intuitionistic logic [26]. Further extensions include formalising Solovay's system GLS and interpretability logics from both semantic [61,79] and proof-theoretic perspectives [51,71]. Finally, building on recent work integrating modal reasoning into higher-order logic [6], we plan to investigate monadic Grz within our framework.

Related Work. Automated theorem proving for modal logics spans specialised Grz-specific systems and general modular frameworks. Dedicated tableau-based provers for Grz are presented in [74] and via the lean tableaux methodology [2,3]. Translational approaches into decidable first-order fragments [27] further enable the use of established FOL provers for Grz. A broader strategy embeds modal systems, including GL and Grz, into higher-order logic to exploit automation features of Isabelle and similar assistants [5,7–10,52,80]; related ideas for handling non-classical reasoning appear in the HOL Light tutorial [53, §20.4].

Within the Prolog ecosystem, decision procedures have been developed for classical and intuitionistic modal logics [38–40,69]. In proof assistants, substantial efforts include the NAMOR library in Agda [22,23], tableau-based automation in Lean [86], and extensive meta-theoretical developments in Isabelle [34–36] and Rocq [28–30,32,44,45,47–49,57,75,76].

Hybrid techniques such as CEGAR-Tableaux enhance satisfiability checking for normal modal logics [42,43], while systems like Goéland [24,25] provide concurrent tableau-based reasoning adaptable to modal settings.

Complementary foundational work clarifies the structural properties underpinning extensible modal decision procedures [56,62–65,70] and explores automated proofs of non-trivial proof-theoretic results [11,12].

We propose our work as a novel and coherent synthesis, in a fully computerised environment, of the mentioned work in: the intrinsic extensibility of decision procedures for modal systems; the mathematical generality provided by rigorous completeness proofs constructed through modular proof strategies; and the potential for further extensions of automated modal reasoning via translations and embeddings, fully formalised.

Acknowledgments. We gratefully acknowledge the constructive feedback provided by three anonymous reviewers on the initial submission.

This work was partially funded by the project SERICS – Security and Rights in the CyberSpace PE0000014, financed within PNRR, M4C2 I.1.3, funded by the European Union - NextGenerationEU (MUR Code: 2022CY2J5S, CUP: D67G22000340001); the Istituto Nazionale di Alta Matematica – INdAM group GNSAGA; the research project "Differential, Algebraic, Complex and Arithmetic Geometry (2025)"; The International Research Network "Logic and Interaction".

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Bibliography

1. Aleksander Wołoszyn, W.: Grzegorczyk logic unlocked. arXiv e-prints arXiv:2505.09836, May 2025. https://doi.org/10.48550/arXiv.2505.09836
2. Beckert, B., Goré, R.: Free-variable tableaux for propositional modal logics. Stud Logica **69**(1), 59–96 (2001). https://doi.org/10.1023/A:1013886427723
3. Beckert, B., Posegga, J.: leanTAP: lean Tableau-based deduction. J. Autom. Reason. **15**(3), 339–358 (1995). https://doi.org/10.1007/BF00881804
4. van Benthem, J., Blackburn, P.: Modal logic: a semantic perspective. In: Handbook of Modal Logic, vol. 3, pp. 1–84. Elsevier (2007)
5. Benzmueller, C., Paulson, L.C.: Multimodal and intuitionistic logics in simple type theory. Logic J. IGPL **18**(6), 881–892 (01 2010). https://doi.org/10.1093/jigpal/jzp080
6. Benzmüller, C.: Universal (meta-)logical reasoning: recent successes. Sci. Comput. Program. **172**, 48–62 (2019). https://doi.org/10.1016/J.SCICO.2018.10.008
7. Benzmüller, C.: Faithful logic embeddings in HOL - a recipe to have it all: deep and shallow, automated and interactive, heavy and light, proofs and counterexamples, meta and object level. CoRR abs/2502.19311 (2025). https://doi.org/10.48550/ARXIV.2502.19311
8. Benzmüller, C., Claus, M., Sultana, N.: Systematic verification of the modal logic cube in Isabelle/HOL. In: Kaliszyk, C., Paskevich, A. (eds.) Proceedings Fourth Workshop on Proof eXchange for Theorem Proving, PxTP 2015, Berlin, Germany, 2–3 August 2015. EPTCS, vol. 186, pp. 27–41 (2015). https://doi.org/10.4204/EPTCS.186.5
9. Benzmüller, C., Woltzenlogel Paleo, B.: Higher-order modal logics: automation and applications. In: Faber, W., Paschke, A. (eds.) Reasoning Web 2015. LNCS, vol. 9203, pp. 32–74. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-21768-0_2
10. Benzmüller, C., Paulson, L.C.: Quantified multimodal logics in simple type theory. Logica Universalis **7**(1), 7–20 (2013). https://doi.org/10.1007/S11787-012-0052-Y
11. Bezhanishvili, N., Ghilardi, S.: Bounded proofs and step frames. In: Galmiche, D., Larchey-Wendling, D. (eds.) TABLEAUX 2013. LNCS (LNAI), vol. 8123, pp. 44–58. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40537-2_6
12. Bezhanishvili, N., Ghilardi, S.: The bounded proof property via step algebras and step frames. Ann. Pure Appl. Log. **165**(12), 1832–1863 (2014). https://doi.org/10.1016/J.APAL.2014.07.005

13. Bilotta, A., Maggesi, M., Perini Brogi, C.: A modular proof of semantic completeness for normal systems beyond the modal cube, formalised in HOLMS. In: Moscardelli, L., Scozzari, F. (eds.) Proceedings of the 26th Italian Conference on Theoretical Computer Science, Pescara, Italy, 10–12 September 2025. CEUR Workshop Proceedings, vol. 4039, pp. 154–162. CEUR-WS.org (2025). https://ceur-ws.org/Vol-4039/paper10.pdf
14. Bilotta, A., Maggesi, M., Perini Brogi, C.: A modular framework for proof-search via formalised modal completeness in HOL light. In: Guerrini, S., König, B. (eds.) 34th EACSL Annual Conference on Computer Science Logic (CSL 2026). Leibniz International Proceedings in Informatics (LIPIcs), vol. 363, pp. 18:1–18:29. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2026). https://doi.org/10.4230/LIPIcs.CSL.2026.18. https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CSL.2026.18
15. Bilotta, A., Maggesi, M., Perini Brogi, C.: Growing HOLMS: a verified automated prover for Grzegorczyk logic in HOL light (extended version) (2026). https://hdl.handle.net/20.500.11771/41542
16. Bilotta, A., Maggesi, M., Perini Brogi, C.: A HOL light library for modal systems (HOLMS) (2026). https://holms-lib.github.io/. Accessed May 2026
17. Bilotta, A., Maggesi, M., Perini Brogi, C., Quartini, L.: Growing HOLMS, a HOL light library for modal systems. In: Porello, D., Vinci, C., Zavatteri, M. (eds.) Short Paper Proceedings of the 6th International Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis, OVERLAY 2024, Bolzano, Italy, 28–29 November 2024. CEUR Workshop Proceedings, vol. 3904, pp. 41–48. CEUR-WS.org (2024). https://ceur-ws.org/Vol-3904/paper5.pdf
18. Blackburn, P., van Benthem, J.F.A.K., Wolter, F. (eds.): Handbook of Modal Logic, Studies in Logic and Practical Reasoning, vol. 3. North-Holland (2007). https://www.sciencedirect.com/bookseries/studies-in-logic-and-practical-reasoning/vol/3/suppl/C
19. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic, Cambridge Tracts in Theoretical Computer Science, vol. 53. Cambridge University Press (2001). https://doi.org/10.1017/CBO9781107050884
20. Boolos, G.: Provability in arithmetic and a schema of Grzegorczyk. Fundamenta Mathematicae **106**(1), 41–45 (1980). http://eudml.org/doc/211104
21. Boolos, G.: The Logic of Provability. Cambridge University Press (1995)
22. Borsetto, R., Zorzi, M.: An agda implementation of the modal logic S4.2: first investigations. In: Moscardelli, L., Scozzari, F. (eds.) Proceedings of the 26th Italian Conference on Theoretical Computer Science, Pescara, Italy, 10–12 September 2025. CEUR Workshop Proceedings, vol. 4039, pp. 163–168. CEUR-WS.org (2025). https://ceur-ws.org/Vol-4039/paper17.pdf
23. Borsetto, R., Zorzi, M.: NAMOR: a new agda library for modal extended sequents. In: Short Paper Proceedings of the 7th International Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis, OVERLAY 2025, Bologna, Italy, 26 October 2025 (2025). https://overlay.uniud.it/workshop/2025/papers/borsetto-zorzi.pdf
24. Cailler, J.: Designing an automated concurrent tableau-based theorem prover for first-order logic. (Conception d'un prouveur automatique de théorèmes concurrent basé sur la méthode des tableaux pour la logique du premier ordre). Ph.D. thesis, University of Montpellier, France (2023). https://tel.archives-ouvertes.fr/tel-04526940

25. Cailler, J., Rosain, J., Delahaye, D., Robillard, S., Bouziane, H.: Goéland: a concurrent tableau-based theorem prover (system description). In: Blanchette, J., Kovács, L., Pattinson, D. (eds.) Automated Reasoning - 11th International Joint Conference, IJCAR 2022, Haifa, Israel, 8–10 August 2022, Proceedings. LNCS, vol. 13385, pp. 359–368. Springer (2022). https://doi.org/10.1007/978-3-031-10769-6_22
26. Chagrov, A.V., Zakharyaschev, M.: Modal Logic, Oxford Logic Guides, vol. 35, Oxford University Press (1997)
27. Demri, S., Goré, R.: An O $((n{\cdot}\log n)^3)$-time transformation from Grz into decidable fragments of classical first-order logic. In: Caferra, R., Salzer, G. (eds.) Automated Deduction in Classical and Non-Classical Logics, Selected Papers. LNCS, vol. 1761, pp. 152–166. Springer (1998). https://doi.org/10.1007/3-540-46508-1_10
28. Doczkal, C., Bard, J.: Completeness and decidability of converse PDL in the constructive type theory of Coq. In: Proceedings of the 7th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2018, pp. 42–52. Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3167088
29. Doczkal, C., Smolka, G.: Constructive formalization of hybrid logic with eventualities. In: Jouannaud, J.P., Shao, Z. (eds.) Certified Programs and Proofs, pp. 5–20. Springer, Heidelberg (2011)
30. Doczkal, C., Smolka, G.: Completeness and decidability results for CTL in constructive type theory. J. Autom. Reason. **56**(3), 343–365 (2016)
31. Dyckhoff, R., Negri, S.: A cut-free sequent system for Grzegorczyk logic, with an application to the Gödel-McKinsey-Tarski embedding. J. Log. Comput. **26**(1), 169–187 (2016). https://doi.org/10.1093/LOGCOM/EXT036
32. Férée, H., van der Giessen, I., van Gool, S., Shillito, I.: Mechanised uniform interpolation for modal logics K, GL, and iSL. In: Benzmüller, C., Heule, M.J.H., Schmidt, R.A. (eds.) Automated Reasoning - 12th International Joint Conference, IJCAR 2024, Nancy, France, 3–6 July 2024, Proceedings, Part II. LNCS, vol. 14740, pp. 43–60. Springer (2024). https://doi.org/10.1007/978-3-031-63501-4_3
33. Fitting, M.: Modal proof theory. In: Handbook of Modal Logic, pp. 85–138. Elsevier (2007)
34. From, A.H.: Formalized soundness and completeness of epistemic logic. In: Silva, A., Wassermann, R., de Queiroz, R. (eds.) WoLLIC 2021. LNCS, vol. 13038, pp. 1–15. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-88853-4_1
35. From, A.H.: An Isabelle/HOL framework for synthetic completeness proofs. In: Proceedings of the 14th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2025, pp. 171–186. Association for Computing Machinery, New York, NY, USA (2025). https://doi.org/10.1145/3703595.3705882
36. From, A.H., Schlichtkrull, A.: Abstract, compositional consistency: Isabelle/HOL locales for completeness à la fitting. In: Forster, Y., Keller, C. (eds.) 16th International Conference on Interactive Theorem Proving, ITP 2025, Reykjavik, Iceland, 28 September–1 October 2025. LIPIcs, vol. 352, pp. 8:1–8:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2025). https://doi.org/10.4230/LIPICS.ITP.2025.8
37. Garg, D., Genovese, V., Negri, S.: Countermodels from sequent calculi in multimodal logics. In: Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science, LICS 2012, Dubrovnik, Croatia, 25–28 June 2012, pp. 315–324. IEEE Computer Society (2012). https://doi.org/10.1109/LICS.2012.42

38. Girlando, M., Lellmann, B., Olivetti, N., Pozzato, G.L., Vitalis, Q.: VINTE: an implementation of internal calculi for Lewis' logics of counterfactual reasoning. In: Schmidt, R.A., Nalon, C. (eds.) TABLEAUX 2017. LNCS (LNAI), vol. 10501, pp. 149–159. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66902-1_9
39. Girlando, M., Lellmann, B., Olivetti, N., Pesce, S., Pozzato, G.L.: Calculi, countermodel generation and theorem prover for strong logics of counterfactual reasoning. J. Log. Comput. (2022). https://doi.org/10.1093/logcom/exab084
40. Girlando, M., Straßburger, L.: MOIN: a nested sequent theorem prover for intuitionistic modal logics (system description). In: Peltier, N., Sofronie-Stokkermans, V. (eds.) IJCAR 2020. LNCS (LNAI), vol. 12167, pp. 398–407. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-51054-1_25
41. Gödel, K.: Eine Interpretation des Intuitionistischen Aussagenkalküls. Ergebnisse eines Mathematischen Kolloquiums, **4**, 39–40 (1933). English translation, with an introductory note by A.S. Troelstra. Kurt Gödel, Collected Works **1**, 296–303 (1986)
42. Goré, R., Kelly, J.: Automated proof search in Gödel-Löb provability logic. In: Abstract, British Logic Colloquium (2007)
43. Goré, R., Kikkert, C.: CEGAR-tableaux: improved modal satisfiability via modal clause-learning and SAT. In: Das, A., Negri, S. (eds.) TABLEAUX 2021. LNCS (LNAI), vol. 12842, pp. 74–91. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-86059-2_5
44. Goré, R., Ramanayake, R., Shillito, I.: Cut-elimination for provability logic by terminating proof-search: formalised and deconstructed using coq. In: Das, A., Negri, S. (eds.) Automated Reasoning with Analytic Tableaux and Related Methods, pp. 299–313. Springer International Publishing, Cham (2021)
45. de Groot, J., Shillito, I., Clouston, R.: Semantical analysis of intuitionistic modal logics between CK and IK. In: 40th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2025, Singapore, 23–26 June 2025, pp. 169–182. IEEE (2025). https://doi.org/10.1109/LICS65433.2025.00020
46. Grzegorczyk, A.: Some relational systems and the associated topological spaces. Fundamenta Mathematicae **60**(3), 223–231 (1967). http://eudml.org/doc/213962
47. Hagemeier, C.: Formalizing intuitionistic epistemic logic in Coq. BSc thesis (2021)
48. Hagemeier, C., Kirst, D.: Constructive and mechanised meta-theory of IEL and similar modal logics. J. Log. Comput. **32**(8), 1585–1610 (2022). https://doi.org/10.1093/LOGCOM/EXAC068
49. Hagemeier, C., Kirst, D.: Constructive and mechanised meta-theory of intuitionistic epistemic logic. In: Artemov, S., Nerode, A. (eds.) LFCS 2022. LNCS, vol. 13137, pp. 90–111. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93100-1_7
50. Hakli, R., Negri, S.: Does the deduction theorem fail for modal logic? Synthese **187**(3), 849–867 (2012)
51. Hakoniemi, T.A., Joosten, J.J.: Labelled tableaux for interpretability logics. In: Liber Amicorum Alberti: A Tribute to Albert Visser (2016)
52. Hardt, M., Smolka, G.: Higher-order syntax and saturation algorithms for hybrid logic. In: Blackburn, P., Bolander, T., Braüner, T., de Paiva, V., Villadsen, J. (eds.) Proceedings of the International Workshop on Hybrid Logic, HyLo@FLoC 2006, Seattle, WA, USA, 11 August 2006. ENTCS, vol. 174, pp. 15–27. Elsevier (2006). https://doi.org/10.1016/J.ENTCS.2006.11.023
53. Harrison, J.: HOL Light tutorial (2017). http://www.cl.cam.ac.uk/~jrh13/hol-light/tutorial.pdf
54. Harrison, J.: The HOL Light system reference (2023). https://www.cl.cam.ac.uk/~jrh13/hol-light/reference.pdf

55. Harrison, J.: The HOL Light Theorem Prover (2025). https://hol-light.github.io/
56. Hustadt, U., Papacchini, F., Nalon, C., Dixon, C.: Model construction for modal clauses. In: International Joint Conference on Automated Reasoning, pp. 3–23. Springer (2024)
57. Kirst, D., Shillito, I.: Completeness of first-order bi-intuitionistic logic. In: Endrullis, J., Schmitz, S. (eds.) 33rd EACSL Annual Conference on Computer Science Logic, CSL 2025, 10–14 February 2025, Amsterdam, Netherlands. LIPIcs, vol. 326, pp. 40:1–40:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2025). https://doi.org/10.4230/LIPICS.CSL.2025.40
58. Maggesi, M., Perini Brogi, C.: A formal proof of modal completeness for provability logic. In: Cohen, L., Kaliszyk, C. (eds.) 12th International Conference on Interactive Theorem Proving (ITP 2021). Leibniz International Proceedings in Informatics (LIPIcs), vol. 193, pp. 26:1–26:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2021). https://doi.org/10.4230/LIPIcs.ITP.2021.26. https://drops.dagstuhl.de/opus/volltexte/2021/13921
59. Maggesi, M., Perini Brogi, C.: Mechanising Gödel-Löb provability logic in HOL light. J. Autom. Reason. **67**(3), 29 (2023). https://doi.org/10.1007/S10817-023-09677-Z
60. McKinsey, J.C.C., Tarski, A.: Some theorems about the sentential calculi of Lewis and Heyting. J. Symbolic Log. **13**(1), 1–15 (1948). http://www.jstor.org/stable/2268135
61. Mikec, L.: Satisfiability verifiers for certain modal logics concerned with provability (2026). https://luka.doublebuffer.net/o/il/. Accessed Feb 2026. https://github.com/luka-mikec/provability_sat
62. Nalon, C.: Efficient theorem-proving for modal logics. In: Ciabattoni, A., Gabelaia, D., Sedlár, I. (eds.) Advances in Modal Logic, AiML 2024, Prague, Czech Republic, 19–23 August 2024, pp. 13–16. College Publications (2024)
63. Nalon, C., Dixon, C., Hustadt, U.: Modal resolution: proofs, layers, and refinements. ACM Trans. Comput. Log. (TOCL) **20**(4), 1–38 (2019)
64. Nalon, C., Hustadt, U., Papacchini, F., Dixon, C.: Local reductions for the modal cube. In: International Joint Conference on Automated Reasoning. pp. 486–505. Springer International Publishing Cham (2022)
65. Nalon, C., Hustadt, U., Papacchini, F., Dixon, C.: Buy one get 14 free: evaluating local reductions for modal logic. In: Pientka, B., Tinelli, C. (eds.) Automated Deduction - CADE 29, pp. 382–400. Springer Nature Switzerland, Cham (2023)
66. Negri, S.: Proof analysis in modal logic. J. Philos. Log. **34**(5), 507–544 (2005)
67. Negri, S.: Proofs and countermodels in non-classical logics. Log. Univers. **8**(1), 25–60 (2014)
68. Negri, S., von Plato, J.: Proof Analysis: A Contribution to Hilbert's Last Problem. Cambridge University Press (2011)
69. Olivetti, N., Pozzato, G.L.: NESCOND: an implementation of nested sequent calculi for conditional logics. In: Demri, S., Kapur, D., Weidenbach, C. (eds.) IJCAR 2014. LNCS (LNAI), vol. 8562, pp. 511–518. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08587-6_39
70. Pattinson, D., Olivetti, N., Nalon, C.: Resolution calculi for non-normal modal logics. In: International Conference on Automated Reasoning with Analytic Tableaux and Related Methods, pp. 322–341. Springer (2023)
71. Perini Brogi, C., Negri, S., Olivetti, N.: Modular sequent calculi for interpretability logics. Rev. Symb. Log. **18**(3), 704–743 (2025). https://doi.org/10.1017/S1755020325100701

72. Poggiolesi, F.: The method of tree-hypersequents for modal propositional logic. In: Makinson, D., Malinowski, J., Wansing, H. (eds.) Towards Mathematical Philosophy. TL, vol. 28, pp. 31–51. Springer, Dordrecht (2009). https://doi.org/10.1007/978-1-4020-9084-4_3
73. Savateev, Y., Shamkanov, D.: Non-well-founded proofs for the Grzegorczyk modal logic. Rev. Symbolic Log. **14**(1), 22–50 (2021). https://doi.org/10.1017/S1755020319000510
74. Schoeber, N.: Automatic Tableau builder for Grzegorczyk logic. Bachelor's Thesis, Faculty of Science and Engineering, University of Groningen (2017)
75. Shillito, I., van der Giessen, I., Goré, R., Iemhoff, R.: A new calculus for intuitionistic strong Löb logic: strong termination and cut-elimination, formalised. In: Ramanayake, R., Urban, J. (eds.) Automated Reasoning with Analytic Tableaux and Related Methods - 32nd International Conference, TABLEAUX 2023, Prague, Czech Republic, 18–21 September 2023, Proceedings. LNCS, vol. 14278, pp. 73–93. Springer (2023). https://doi.org/10.1007/978-3-031-43513-3_5
76. Shillito, I., Kirst, D.: A mechanised and constructive reverse analysis of soundness and completeness of bi-intuitionistic logic. In: Timany, A., Traytel, D., Pientka, B., Blazy, S. (eds.) Proceedings of the 13th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2024, London, UK, 15–16 January 2024, pp. 218–229. ACM (2024). https://doi.org/10.1145/3636501.3636957
77. Simpson, A.K.: The Proof Theory and Semantics of Intuitionistic Modal Logic. Ph.D. thesis, University of Edinburgh. College of Science and Engineering. School of Informatics (1994)
78. Smorynski, C.: The incompleteness theorems. In: Barwise, J. (ed.) Handbook of Mathematical Logic, pp. 821–865. North-Holland (1977)
79. Solovay, R.M.: Provability interpretations of modal logic. Israel J. Math. **25**(3–4), 287–304 (1976)
80. Steen, A., Sutcliffe, G., Benzmüller, C.: Solving quantified modal logic problems by translation to classical logics. J. Log. Comput., exaf006 (2025)
81. Strachey, C.: Fundamental concepts in programming languages. Higher-Order Symbolic Comput. **13**, 11–49 (2000)
82. Troelstra, A.S., Schwichtenberg, H.: Basic Proof Theory, Cambridge Tracts in Theoretical Computer Science, 2nd edn., vol. 43. Cambridge University Press (2000)
83. Tseitin, G.S.: On the complexity of derivation in propositional calculus, pp. 466–483. Springer, Heidelberg (1983). https://doi.org/10.1007/978-3-642-81955-1_28
84. Verbrugge, R.L.: Provability logic. In: Zalta, E.N., Nodelman, U. (eds.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Summer 2024 edn. (2024)
85. Viganò, L.: Labelled Non-Classical Logics. Springer Science & Business Media (2013)
86. Wu, M., Goré, R.: Verified decision procedures for modal logics. In: Harrison, J., O'Leary, J., Tolmach, A. (eds.) 10th International Conference on Interactive Theorem Proving (ITP 2019). Leibniz International Proceedings in Informatics (LIPIcs), vol. 141, pp. 31:1–31:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2019). https://doi.org/10.4230/LIPIcs.ITP.2019.31. https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ITP.2019.31