Cut-Elimination for Provability Logic by Terminating Proof-Search: Formalised and Deconstructed Using Coq

Rajeev Goré[3], Revantha Ramanayake[2], and Ian Shillito[1](✉)

[1] Australian National University, Canberra, ACT, Australia
ian.shillito@anu.edu.au
[2] University of Groningen, Groningen, The Netherlands
d.r.s.ramanayake@rug.nl
[3] Technische Universität Wien, Vienna, Austria

Abstract. Recently, Brighton gave another cut-admissibility proof for the standard set-based sequent calculus GLS for modal provability logic GL. One of the two induction measures that Brighton uses is novel: the maximum height of regress trees in an auxiliary calculus called RGL. Tautology elimination is established rather than direct cut-admissibility, and at some points the input derivation appears to be ignored in favour of a derivation obtained by backward proof-search. By formalising the GLS calculus and the proofs in Coq, we show that: (1) the use of the novel measure is problematic under the usual interpretation of the Gentzen comma as set union, and a multiset-based sequent calculus provides a more natural formulation; (2) the detour through tautology elimination is unnecessary; and (3) we can use the same induction argument without regress trees to obtain a direct proof of cut-admissibility that is faithful to the input derivation.

Keywords: Provability logic · Cut admissibility · Interactive theorem proving · Proof theory

1 Introduction

Propositional modal provability logics extend the basic normal modal logic K with axioms which interpret the $\Box$ connective as the mathematical notion of being "provable" in Peano Arithmetic [1,16]. There are several variants with characteristic axioms named after Gödel, Löb and Grzegorczyk:

Name	Characteristic Axiom
GL	$\Box(\Box p \to p) \to \Box p$
Go	$\Box(\Box(p \to \Box p) \to p) \to \Box p$
Grz	$\Box(\Box(p \to \Box p) \to p) \to p$

While the "provability" interpretation is now well-understood, the proof-theory of these logics is intricate and somewhat controversial as we explain next.

© Springer Nature Switzerland AG 2021
A. Das and S. Negri (Eds.): TABLEAUX 2021, LNAI 12842, pp. 299–313, 2021.
https://doi.org/10.1007/978-3-030-86059-2_18

Following Gentzen [5,6], the literature abounds with proofs of cut-admissibility for various sequent calculi based on the size of the cut-formula and the height of the premise derivations. But these measures looked, at first sight, inadequate for proving cut-elimination for the standard set-based sequent calculus GLS for provability logic GL so Valentini introduced a third novel measure called width, and showed that cut-elimination for GLS could be obtained via a triple induction over size, height and width [17].

Controversy arose when it was (erroneously) claimed that Valentini's proofs contained a gap and various authors provided alternative proofs of cut-elimination in response [2,10–12,14]. The question was resolved in Valentini's favour [7], with all proofs later verified using an interactive theorem prover Isabelle/HOL [4].

The cut-elimination proof for the logic Go (due to Goré and Ramanayake [8] via a deeper analysis of the structure of derivations, and subsequently by Savateev and Shamkanov [15] via non-well founded-proofs) is even more intricate. The proof-theory of provability logics can therefore be described as complex.

Recently, Brighton [3] provided yet another proof of cut-admissibility for GLS which is significantly simpler than any of the existing proofs of cut-admissibility in the literature. It uses a double induction with the traditional size of the cut-formula as primary measure. The secondary measure is called the "maximum height of regress trees" and it is a novel measure defined using a backward proof-search procedure for GLS called RGL, based on regress trees/regressants.

Backward proof-search can often be employed to obtain cut-free completeness with respect to the Kripke semantics of a logic. However, cut-elimination is *not* a result directly obtained by the use of backward proof-search. For this reason Brighton's method is intriguing from a structural proof theoretic perspective. Even more so because, from a tableaux perspective, the RGL calculus is nothing but the backward proof-search decision procedure for GL that is well-known to be cut-free complete with respect to the Kripke semantics of GL. Unfortunately, Brighton's arguments is clouded by various issues that become apparent when studying them in detail.

We first explain why Brighton's use of a set-based sequent calculus leads to confusion, and explain how this can be clarified using multisets. We then show that the special calculus RGL on regress trees can be replaced by a standard proof-search procedure PSGLS on GLS itself. Putting this all together, we replace Brighton's detour through tautology elimination [9] with a direct proof of cut-admissibility for GL making use of the maximum height of a derivation (the existence of the latter follows from the termination of backward proof-search). Noting that Brighton's proof seems to ignore the structure of the given cut-free derivations of the premises, and since such a shortcoming undermines cut-elimination as a procedure that manipulates the given derivations to produce a cut-free derivation, we take particular care to highlight the local nature of our transformations. All of our claims have been formally verified in the interactive theorem prover Coq (https://github.com/ianshil/CE_GLS.git).

2 Various Issues with the Method Used by Brighton

Although Brighton's work is extremely appealing, we have already mentioned that the argument and the proof technique supporting it require further clarifications. Let us exhibit the two main elements that appeared through the formalisation process to be responsible for this unclarity.

First, as the sequents that are used are based on sets, the rule for implication on the right, presented below on the left, is just a notation for the rule on the right where the comma is interpreted as set union.

$$\frac{A, X \Rightarrow Y, B}{X \Rightarrow Y, A \to B} \qquad \frac{\{A\} \cup X \Rightarrow Y \cup \{B\}}{X \Rightarrow Y \cup \{A \to B\}}\ (\to\mathrm{R})$$

However, it is well-known that $(\to\mathrm{R})$ contains an implicit contraction [7]. As a consequence, $(\to\mathrm{R})$ could be reapplied as many times as one wants above $\Rightarrow p \to q$ on the formula $p \to q$. That implies the existence of derivations of all heights for this sequent, as shown below.

$$\cfrac{\cfrac{\{p\} \cup \emptyset \Rightarrow \emptyset \cup \{q\} \cup \{p \to q\}}{\{p\} \cup \emptyset \Rightarrow \emptyset \cup \{q\} \cup \{p \to q\}}\ (\to\mathrm{R})}{\emptyset \Rightarrow \emptyset \cup \{p \to q\}}\ (\to\mathrm{R})$$

Brighton's argument requires (and proves) that all sequents have a derivation of maximum height - this would contradict our observation above. For his argument to hold, it must therefore be the case that Brighton is not using the usual interpretation for the rules $(\to\mathrm{R})$ and $(\to\mathrm{L})$.

The only reasonable option seems to be that Brighton intends for the comma to be interpreted as *disjoint union*. This amounts to the following rule.

$$\frac{\{A\} \cup X \Rightarrow (Y \setminus \{A \to B\}) \cup \{B\}}{X \Rightarrow Y \cup \{A \to B\}}\ (\to\mathrm{R}_{\mathrm{Dis}})$$

If that was the case, a proof that the calculus is complete for GL under this interpretation is required. Moreover, further issues arise with this interpretation.

For example, it is not true in general that the premise of the sequent $\Gamma \Rightarrow \Delta, B \to C$ via the rule $(\to\mathrm{R}_{\mathrm{Dis}})$ is $B, \Gamma \Rightarrow \Delta, C$ (Case 2 of Theorem 1 of Brighton's article). Indeed, if $B \to C \in \Delta$ then $B, \Gamma \Rightarrow \Delta, C$ and $B, \Gamma \Rightarrow (\Delta \setminus \{B \to C\}), C$ would be different. This issue seems repairable. However the situation is undesirable given the sensitivity of structural proof theory to small syntactic details, and especially given the history of cut-elimination for GL.

Second, Brighton provides an unusual argument for the admissibility of cut. In order to obtain a proof of the latter, Brighton proves a result equivalent to it in the case of classical calculi: tautology elimination. More precisely, this lemma has the following shape: if $A \to A, X \Rightarrow Y$ is provable then so is $X \Rightarrow Y$. On inspection, it is clear that a procedure for tautology elimination can easily be turned into a procedure for cut-elimination, and *vice versa*. Given the proximity between these results, arguing for the admissibility of cut by proving tautology elimination seems to be an unnecessary detour.

3 Preliminaries

Let $\mathbb{V} = \{p, q, r \dots\}$ be an infinite set of propositional variables. Modal formulae are defined by the following grammar.

$$A ::= p \in \mathbb{V} \mid \bot \mid A \to A \mid \Box A$$

We use a minimal set of connectives since it is well-known that the other connectives can be defined from these.

We define the *size* of a formula by the number of symbols it contains. We say that a formula A is a *boxed formula* if it has $\Box$ as its main connective. A boxed multiset contains only boxed formulae. For a set $X = \{A_1, \dots, A_n\}$, define $\boxtimes X = \{A_1, \Box A_1, \dots, A_n, \Box A_n\}$. We denote the set of subformulae of a formula A by $\mathrm{Sub}(A)$. We abuse the notation to designate the set of subformulae of all formulae in the set X by $\mathrm{Sub}(X)$. In what follows we use the letters $A, B, C, \dots$ for formulae and $X, Y, Z, \dots$ for multisets of formulae.

The Hilbert calculus for the basic normal modal logic K extends a Hilbert-calculus for classical propositional logic with the axiom $\Box(p \to q) \to (\Box p \to \Box q)$ and the inference rule of necessitation: from A infer $\Box A$. Gödel-Löb logic GL is obtained by the addition of the axiom $\Box(\Box p \to p) \to \Box p$ to K.

A *sequent* is a pair of multisets of formulae, denoted $X \Rightarrow Y$. For multisets X and Y, the multiset sum $X \uplus Y$ is the multiset whose multiplicity (at each formula) is a sum of the multiplicities of X and Y. We write X, Y to mean $X \uplus Y$. For a formula A, we write A, X and X, A to mean $\{A\} \uplus X$.

A *sequent calculus* consists of a finite set of *sequent rule schemas*. Each rule schema consists of a conclusion sequent and some number of premise sequents. If a rule schema has no premise sequents, then it is called an initial sequent. The conclusion and premises are built in the usual way from propositional-variables, formula-variables and multiset-variables. A *rule instance* is obtained by uniformly instantiating every variable in the rule schema with a concrete object of that type. This is the standard definition from structural proof theory.

Definition 1 (Derivation/Proof). *A* derivation *of a sequent s in the sequent calculus C is a finite tree of sequents such that (i) the root node is s; and (ii) each interior node and its direct children are the conclusion and premise(s) of a rule instance in C.*

A proof *is a derivation where every leaf is an instance of an initial sequent.*

In what follows, it should be clear from context whether the word "proof" refers to the object defined in Definition 1, or to the meta-level notion. We say that a sequent is *provable* in C if it has a proof in C.

Definition 2 (Height). *The* height *of a derivation δ, noted $h(\delta)$, is the maximum number of nodes on a path from root to leaf.*

The sequent calculus GLS is given in Fig. 1.

$$\frac{}{p, X \Rightarrow Y, p}\ (\text{IdP}) \qquad \frac{}{\bot, X \Rightarrow Y}\ (\bot\text{L})$$

$$\frac{X \Rightarrow Y, A \qquad B, X \Rightarrow Y}{A \to B, X \Rightarrow Y}\ (\to\text{L}) \qquad \frac{A, X \Rightarrow Y, B}{X \Rightarrow Y, A \to B}\ (\to\text{R})$$

$$\frac{\boxtimes X, \Box B \Rightarrow B}{W, \Box X \Rightarrow \Box B, \Box Y, Z}\ (\text{GLR})$$

Fig. 1. The sequent calculus GLS. Here, W and Z do not contain any boxed formulae.

In a rule instance of ($\to$L) or ($\to$R), the formula instantiating the featured $A \to B$ is the *principal formula* of that instance. In (IdP), a propositional variable instantiating either featured occurrence of p is principal. In a rule instance of (GLR), the formula $\Box B$ is called the *diagonal formula* [13].

Example 1. The following are examples of derivations in GLS. Note that while the first and second examples are derivations, the third is a proof.

$$p \Rightarrow q \to r \qquad \frac{p, q \Rightarrow r}{p \Rightarrow q \to r}\ (\to\text{R}) \qquad \cfrac{\cfrac{}{\Box p, p, \Box p \Rightarrow p}\ (\text{IdP})}{\Box p \Rightarrow \Box p}\ (\text{GLR})$$

Example 2. A special example of derivation in GLS is the following:

$$\cfrac{\cfrac{\Box A \to A, \Box(\Box A \to A), A, A, \Box A, \Box A, \Box A \Rightarrow A}{\Box(\Box A \to A), A, \Box A, \Box A \Rightarrow A, \Box A}\ (\text{GLR}) \qquad \Box(\Box A \to A), A, A, \Box A, \Box A \Rightarrow A}{\Box A \to A, \Box(\Box A \to A), A, \Box A, \Box A \Rightarrow A}\ (\to\text{L})$$

By noticing the identity modulo formula multiplicities between the topmost and the lowest sequents, it appears that the sequence of application of rules in the above could be iterated indefinitely on the topmost sequent.

Finally, we consider the additive cut rule.

$$\frac{X \Rightarrow Y, A \qquad A, X \Rightarrow Y}{X \Rightarrow Y}\ (\text{cut})$$

In the above, we call A the *cut-formula*. It easily follows that GLS + (cut) is a sequent calculus for GL [13].

Theorem 1. *For all A we have: $A \in$ GL iff $\Rightarrow A$ is provable in GLS + (cut).*

4 Properties Of GLS

We need some lemmas that are commonly used in proof theory. Straightforward inductions on the structure of formulae or derivations are used to prove them.

Lemma 1. *For all X, Y and A, the sequent $A, X \Rightarrow Y, A$ has a proof.*

Lemma 2 (Height-preserving admissibility of weakening). *For all X, Y, A and B:*

(i) If $X \Rightarrow Y$ has a proof π in GLS*, then $X \Rightarrow Y, A$ has a proof π_0 in* GLS *such that $h(\pi_0) \leq h(\pi)$.*
(ii) If $X \Rightarrow Y$ has a proof π in GLS*, then $A, X \Rightarrow Y$ has a proof π_0 in* GLS *such that $h(\pi_0) \leq h(\pi)$.*

Lemma 3 (Height-preserving invertibility of the implication rules). *For all X, Y, A and B:*

(i) If $A \rightarrow B, X \Rightarrow Y$ has a proof π in GLS*, then $X \Rightarrow Y, A$ and $B, X \Rightarrow Y$ have proofs π_0 and π_1 in* GLS *such that $h(\pi_0) \leq h(\pi)$ and $h(\pi_1) \leq h(\pi)$.*
(ii) If $X \Rightarrow Y, A \rightarrow B$ has a proof π in GLS*, then $A, X \Rightarrow Y, B$ has a proof π_0 in* GLS *such that $h(\pi_0) \leq h(\pi)$.*

Lemma 4 (Height-preserving admissibility of contraction). *For all X, Y, A and B:*

(i) If $X \Rightarrow Y, A, A$ has a proof π in GLS*, then $X \Rightarrow Y, A$ has a proof π_0 in* GLS *such that $h(\pi_0) \leq h(\pi)$.*
(ii) If $A, A, X \Rightarrow Y$ has a proof π in GLS *of height n, then $A, X \Rightarrow Y$ has a proof π_0 in* GLS *such that $h(\pi_0) \leq h(\pi)$.*

In the following section we will introduce a proof-search procedure for GLS which terminates. This will allow us to define the maximum height of a derivation of a sequent with respect to this procedure. Later on this will constitute the secondary induction measure in the proof of admissibility of cut.

5 PSGLS: A Terminating Proof-Search

Given a sequent calculus C, one can define a proof-search procedure on C by imposing further constraints on the applicability of the rules of C. This procedure captures a subset of the set of all derivations of C, i.e. those which are built using the restricted version of the rules of C. Consequently, a proof-search procedure can be identified with the calculus PSC consisting of these restricted rules of C, under the condition that PSC allows to decide the provability of sequents in C.

The sequent calculus PSGLS restricts the rules of GLS in the following way.

1. An additional identity rule (IdB), derivable in GLS as shown in Lemma 1, is introduced.

$$\frac{}{\Box A, X \Rightarrow Y, \Box A}\ (\mathrm{IdB})$$

2. The conclusion of the rule (GLR) is not permitted to be an instance of either (IdP) or ($\bot$L) or (IdB). This restriction ensures that repetitions (even in the weak sense of Example 2) of a sequent along a branch are forbidden.

By inspection, a sequent is provable in PSGLS if and only if it is provable in GLS. The remainder of this section is devoted to showing that each sequent has a derivation of maximum height in PSGLS (something that does not hold of GLS). This crucial result is not thoroughly proved in Brighton's work.

It is easy to prove that if there is a measure that decreases, given a well-founded order, upwards through the rules of PSGLS, then each sequent has a derivation of maximum height in PSGLS. We need the following definition.

Definition 3. *For a sequent $X \Rightarrow Y$:*

1. *Let $\iota(X \Rightarrow Y)$ be the* number of occurrences of "$\rightarrow$" *in $X \Rightarrow Y$.*
2. *Let $\beta(X \Rightarrow Y)$ be the* usable boxes *of $X \Rightarrow Y$ where:*

$$\beta(X \Rightarrow Y) := \{\Box A \mid \Box A \in Sub(X \cup Y)\} \setminus \{\Box A \mid \Box A \in X\}$$

3. *The tuple $(Card(\beta(X \Rightarrow Y)), \iota(X \Rightarrow Y))$, where $Card(U)$ is the cardinality of the set U, is denoted $\Theta(X \Rightarrow Y)$.*

The notion of usable boxes of a sequent $X \Rightarrow Y$ is the set of boxed formulae of $X \Rightarrow Y$ minus the boxed formulae in X. Intuitively, this notion captures the set of boxed formulae of a sequent s which might be the diagonal formula of an instance of (GLR) in a derivation of s in PSGLS.

We proceed to prove that the measure Θ decreases on the usual component-wise ordering on n-tuples, which is well-known to be well-founded, upwards through the rules of PSGLS.

Lemma 5. *Let s_0 and $s_1, ..., s_n$ be sequents. If there is an instance of a rule r of* PSGLS *of the following form, then $\Theta(s_i) < \Theta(s_0)$ for $1 \leq i \leq n$.*

$$\frac{s_1 \quad \dots \quad s_n}{s_0}\ r$$

Proof. We reason by case analysis on r:

1. If r is (IdP) or (IdB) or ($\bot$L), then we are done as there is no premise.
2. If r is ($\rightarrow$R), then it must have the following form.

$$\frac{X, A \Rightarrow Y, B}{X \Rightarrow Y, A \rightarrow B}\ (\rightarrow\mathrm{R})$$

   Then we distinguish two cases. If A is boxed, then $\{\Box B \mid \Box B \in X\} \subseteq \{\Box B \mid \Box B \in X \cup \{A\}\}$. As a consequence, we have that $\beta(X, A \Rightarrow Y, B) \subseteq \beta(X \Rightarrow Y, A \rightarrow B)$ hence $\mathrm{Card}(\beta(X, A \Rightarrow Y, B)) \leq \mathrm{Card}(\beta(X \Rightarrow Y, A \rightarrow B))$. If $\mathrm{Card}(\beta(X, A \Rightarrow Y, B)) < \mathrm{Card}(\beta(X \Rightarrow Y, A \rightarrow B))$ then we are done. If $\mathrm{Card}(\beta(X, A \Rightarrow Y, B)) = \mathrm{Card}(\beta(X \Rightarrow Y, A \rightarrow B))$, then we can see that $\iota(X, A \Rightarrow Y, B) = \iota(X \Rightarrow Y, A \rightarrow B) - 1$ hence $\Theta(X, A \Rightarrow Y, B) < \Theta(X \Rightarrow Y, A \rightarrow B)$. If A is not boxed, then obviously we get that $\mathrm{Card}(\beta(X, A \Rightarrow Y, B)) = \mathrm{Card}(\beta(X \Rightarrow Y, A \rightarrow B))$ but also $\iota(X, A \Rightarrow Y, B) = \iota(X \Rightarrow Y, A \rightarrow B) - 1$ hence $\Theta(X, A \Rightarrow Y, B) < \Theta(X \Rightarrow Y, A \rightarrow B)$.

3. If r is ($\rightarrow$L), then it must have the following form.

$$\frac{X \Rightarrow Y, A \qquad B, X \Rightarrow Y}{A \rightarrow B, X \Rightarrow Y}\ (\rightarrow\mathrm{L})$$

   We can easily establish that $\Theta(X \Rightarrow Y, A) < \Theta(A \rightarrow B, X \Rightarrow Y)$ as one implication symbol is deleted while the cardinality of usable boxes stays the same. To prove that $\Theta(B, X \Rightarrow Y) < \Theta(A \rightarrow B, X \Rightarrow Y)$ we reason as in (2).
4. If r is (GLR) then it must have the following form.

$$\frac{\boxtimes X, \Box B \Rightarrow B}{W, \Box X \Rightarrow \Box B, \Box Y, Z}\ (\mathrm{GLR})$$

   Clearly, we have that $\{\Box A \mid \Box A \in \mathrm{Sub}(\boxtimes X \cup \{\Box B\} \cup \{B\})\} \subseteq \{\Box A \mid \Box A \in \mathrm{Sub}(W \cup \Box X \cup \{\Box B\} \cup \Box Y \cup Z)\}$. Also, given that we consider a derivation in PSGLS, we can note that (IdB) is not applicable on $W, \Box X \Rightarrow \Box B, \Box Y, Z$ by assumption, hence $\Box B \notin \Box X$. Consequently, we get $\{\Box A \mid \Box A \in W \cup \Box X\} \subset \{\Box A \mid \Box A \in \boxtimes X \cup \{\Box B\}\}$. An easy set-theoretic argument leads to $\beta(\boxtimes X, \Box B \Rightarrow B) \subset \beta(W, \Box X \Rightarrow \Box B, \Box Y, Z)$. As a consequence we obtain $\mathrm{Card}(\beta(\boxtimes X, \Box B \Rightarrow B)) < \mathrm{Card}(\beta(W, \Box X \Rightarrow \Box B, \Box Y, Z))$, hence $\Theta(\boxtimes X, \Box B \Rightarrow B) < \Theta(W, \Box X \Rightarrow \Box B, \Box Y, Z)$.

Q.E.D.

The previous lemma implies the existence of a derivation in PSGLS of maximum height for all sequent. We present the formalisation of that theorem, called `PSGLS_termin` in Coq:

```
Theorem PSGLS_termin :
 forall (s : rel (list (MPropF V))),
 existsT2 (DMax: derrec PSGLS_rules (fun _ => True) s),
 (is_mhd DMax).
```

We first universally quantify (`forall`) over the sequent `s`: a *pair* (`rel`) of *lists* (`list`) of *formulae* (`MPropF V`) obtained from the set $\mathbb{V}$ (`V`). Note that while our pen-and-paper proof defines sequents using multisets, our formalisation defines them using lists. The equivalence of these approaches is witnessed by our proof of the derivability of exchange given in our formalisation. Second, we specify that there exists (`existsT2`) an inhabitant `DMax` of the type `derrec PSGLS_rules ( fun _ => True) s`. This is the type of all derivations of s in PSGLS. The ternary function `derrec` outputs a type of finite trees, i.e. derivations in our case, taking as input a set of rules (`PSGLS_rules`), a function describing the set of allowed leaves (`(fun _ => True)`), and the sequent at the root `s`. Third, we state that `DMax` satisfies the property `is_mhd`: it is a derivation of maximum height for the sequent `s`. This formalisation thus corresponds to the following:

Theorem 2. *Every sequent s has a derivation in* PSGLS *of maximum height.*

Proof. We reason by strong induction on the ordered pair $\Theta(s)$. As the applicability of the rules of PSGLS is decidable, we distinguish two cases:

(I) No PSGLS rule is applicable to s. Then the derivation of maximum height sought after is simply the derivation constituted of s solely, which is the only derivation for s.

(II) Some PSGLS rule is applicable to s. Either only initial rules are applicable, in which case the derivation of maximum height sought after is simply the derivation of height 1 constituted of the application of the applicable initial rule to s. Or, some other rules than the initial rules are applicable. Then consider the finite list $Prems(s)$ of all sequents s_0 such that there is an application of a PSGLS rule r with s as conclusion of r and s_0 as premise of r. By Lemma 5 we know that every element s_0 in the list $Prem(s)$ is such that $\Theta(s_0) < \Theta(s)$. Consequently, the induction hypothesis allows us to consider the derivation of maximum height of all the sequents in $Prem(s)$. As $Prem(s)$ is finite, there must be an element s_{max} of $Prem(s)$ such that its derivation of maximum height is higher or of same height than the derivation of maximum height of all sequents in $Prem(s)$. It thus suffices to pick that s_{max}, use its derivation of maximum height, and apply the appropriate rule to obtain s as a conclusion: this is by choice the derivation of maximum height of s.

Q.E.D.

As the previous lemma implies the existence of a derivation δ of maximum height in PSGLS for any sequent s, we are entitled to let $\text{mhd}(s)$ denote the height of δ. Similarly to Brighton, we later use $\text{mhd}(s)$ as the secondary induction measure used in the proof of admissibility of cut.

Before proving the only property we need from $\text{mhd}(s)$, let us interpret the previous lemma from the point of view of the proof-search procedure underlying PSGLS. The existence of a derivation of maximum height for each sequent in PSGLS shows that in the backward application of rules of PSGLS on a sequent, i.e. the carrying of the proof-search procedure, a halting point has to be encountered. As a consequence, the proof-search procedure is *terminating*.

While this is the essence of the content of the previous lemma, we effectively only use the fact that $\text{mhd}(s)$ decreases upwards in the rules of PSGLS.

Lemma 6. *If r is a rule instance from* PSGLS *with conclusion s_0 and s_1 as one of the premises, then $mhd(s_1) < mhd(s_0)$.*

Proof. Suppose that $\text{mhd}(s_1) \geq \text{mhd}(s_0)$. Let δ_0 and δ_1 be the derivations of, respectively, s_0 and s_1 witnessing Theorem 2. Then the following δ_2 is derivation of s_0 of height $\text{mhd}(s_1) + 1$.

$$\dfrac{\begin{matrix}\vdots\, d \\ s_1\end{matrix} \qquad \cdots}{s_0}\, r$$

Because of the maximality of δ_0, we get that the height of δ_0 is greater than the height of δ_2, i.e. $\text{mhd}(s_1) + 1 \leq \text{mhd}(s_0)$. As our initial assumption implies that $\text{mhd}(s_1) + 1 > \text{mhd}(s_0)$, we reached a contradiction. Q.E.D.

Coq is constructive, so how does it allow a proof by contradiction? It can do a proof by contradiction (without having to introduce classical axioms) when dealing with an expression of the decidable fragment. Here, $\mathrm{mhd}(s_1) < \mathrm{mhd}(s_0)$ can be decided because mhd is computable.

6 Cut-Elimination for GLS

We are ready to state and prove our main theorem. It is formalised in Coq in the following way:

```
Theorem GLS_cut_adm : forall A X0 X1 Y0 Y1,
 (derrec GLS_rules (fun _ =>False) (X0++X1,Y0++A::Y1))->
 (derrec GLS_rules (fun _ =>False) (X0++A::X1,Y0++Y1))->
 (derrec GLS_rules (fun _ =>False) (X0++X1,Y0++Y1)).
```

The usual operations on lists "append" and "cons" are respectively represented by `++` and `::`. Sequents are *pairs of lists*, so e.g. `(X0++X1,Y0++Y1)` corresponds to $X_0, X_1 \Rightarrow Y_0, Y_1$. This time `derrec` takes the set of rules `GLS_rules` and the characteristic function `(fun _ => False)` as arguments. So, each line states the existence of a *proof* in GLS. The additive cut rule is formalised in Coq as follows.

$$\frac{\texttt{(X}_0\texttt{++X}_1\texttt{,Y}_0\texttt{++A::Y}_1\texttt{)} \qquad \texttt{(X}_0\texttt{++A::X}_1\texttt{,Y}_0\texttt{++Y}_1\texttt{)}}{\texttt{(X}_0\texttt{++X}_1\texttt{,Y}_0\texttt{++Y}_1\texttt{)}}$$

It is now clear that this statement formalises the following theorem:

Theorem 3. *The additive cut rule is admissible in* GLS.

Proof. Let d_1 (with last rule r_1) and d_2 (with last rule r_2) be proofs in GLS of $X \Rightarrow Y, A$ and $A, X \Rightarrow Y$ respectively, as shown below.

$$\frac{d_1}{X \Rightarrow Y, A}\, r_1 \qquad \frac{d_2}{A, X \Rightarrow Y}\, r_2$$

It suffices to show that there is a proof in GLS of $X \Rightarrow Y$. We reason by strong primary induction (PI) on the size of the cut-formula A, giving the primary inductive hypothesis (PIH), and strong secondary induction (SI) on $\mathrm{mhd}(s)$ of the conclusion of a cut, giving the secondary inductive hypothesis (SIH).

There are five cases to consider for r_1: one for each rule in GLS. We separate them by using Roman numerals. The SIH is invoked in all of the following cases: (III-a), (III-b-1), (III-b-2), (IV) and (V-a-2).

(I) $\mathbf{r_1}$ =(IdP): If A is not principal in r_1, then the latter must have the following form.

$$\frac{}{X_0, p \Rightarrow Y_0, p, A}\,(\mathrm{IdP})$$

where $X_0, p = X$ and $Y_0, p = Y$. Thus, we have that the sequent $X \Rightarrow Y$ is of the form $X_0, p \Rightarrow Y_0, p$, and is an instance of an initial sequent. So we are done.

If A principal in r_1, i.e. $A = p$, then $X \Rightarrow Y$ is of the form $X_0, p \Rightarrow Y$. Thus, the conclusion of r_2 is of the form $X_0, p, p \Rightarrow Y$. We can consequently apply Lemma 4 (ii) to obtain a proof of $X_0, p \Rightarrow Y$.

(II) $\mathbf{r_1}$ =($\bot$L): Then r_1 must have the following form.

$$\frac{}{X_0, \bot \Rightarrow Y, A}\ (\bot\mathrm{L})$$

where $X_0, \bot = X$. Thus, we have that the sequent $X \Rightarrow Y$ is of the form $X_0, \bot \Rightarrow Y$, and is an instance of an initial sequent. So we are done.

(III) $\mathbf{r_1}$ =($\to$ R): We distinguish two cases.

(III-a) If A is not principal in r_1, then the latter must have the following form.

$$\frac{X, B \Rightarrow Y_0, C, A}{X \Rightarrow Y_0, B \to C, A}\ (\to\mathrm{R})$$

where $Y_0, B \to C = Y$. Thus, we have that the sequent $X \Rightarrow Y$ and $A, X \Rightarrow Y$ are respectively of the form $X \Rightarrow Y_0, B \to C$ and $A, X \Rightarrow Y_0, B \to C$. We can apply Lemma 3 (ii) on the proof of the latter to get a proof of $A, X, B \Rightarrow Y_0, C$. Thus proceed as follows.

$$\dfrac{\begin{array}{c} X, B \Rightarrow Y_0, C, A \qquad A, X, B \Rightarrow Y_0, C \\ \hdashline X, B \Rightarrow Y_0, C \end{array}\ \mathrm{SIH}}{X \Rightarrow Y_0, B \to C}\ (\to\mathrm{R})$$

Note that the use of SIH is justified here since the last rule in this proof is an instance of ($\to$R) in PSGLS and hence $\mathrm{mhd}(X, B \Rightarrow Y_0, C) < \mathrm{mhd}(X \Rightarrow Y_0, B \to C)$ by Lemma 6.

(III-b) If A principal in r_1, i.e. $A = B \to C$, then r_1 must have the following form.

$$\frac{B, X \Rightarrow Y, C}{X \Rightarrow Y, B \to C}\ (\to\mathrm{R})$$

The conclusion of r_2 must be of the form $B \to C, X \Rightarrow Y$. In that case, we distinguish two further cases. In the first case, $B \to C$ is principal in r_2. Consequently the latter must have the following form.

$$\frac{X \Rightarrow Y, B \qquad C, X \Rightarrow Y}{B \to C, X \Rightarrow Y}\ (\to\mathrm{L})$$

Proceed as follows.

$$\begin{array}{c} X \Rightarrow Y, B \qquad \begin{array}{c} B, X \Rightarrow Y, C \qquad \begin{array}{c} C, X \Rightarrow Y \\ \hdashline C, B, X \Rightarrow Y \end{array}\ \text{Lem.2 (ii)} \\ \hdashline B, X \Rightarrow Y \end{array}\ \mathrm{PIH} \\ \hdashline X \Rightarrow Y \end{array}\ \mathrm{PIH}$$

In the second case, $B \to C$ is not principal in r_2. In the cases where r_2 is one of (IdP) and ($\bot$L) proceed respectively as in **(I)** and **(II)** when the cut-formula is not principal in the rule considered. We are left with the cases where r_2 is one of ($\to$R), ($\to$L) and (GLR).

(III-b-1) If r_2 is ($\to$R) then it must have the following form.

$$\dfrac{B \to C, D, X \Rightarrow Y_0, E}{B \to C, X \Rightarrow Y_0, D \to E}\ (\to\mathrm{R})$$

where $Y_0, D \to E = Y$. In that case, note that the provable sequent $X \Rightarrow Y, B \to C$ is of the form $X \Rightarrow Y_0, D \to E, B \to C$. We can use Lemma 3 (ii) on the proof of the latter to get a proof of $D, X \Rightarrow Y_0, E, B \to C$. Proceed as follows.

$$\dfrac{\dfrac{D, X \Rightarrow Y_0, E, B \to C \qquad B \to C, D, X \Rightarrow Y_0, E}{D, X \Rightarrow Y_0, E}\ \mathrm{SIH}}{X \Rightarrow Y_0, D \to E}\ (\to\mathrm{R})$$

Note that the use of SIH is justified here as the last rule in this proof is effectively an instance of ($\to$R) in PSGLS, hence $\mathrm{mhd}(X, D \Rightarrow Y_0, E) < \mathrm{mhd}(X \Rightarrow Y_0, D \to E)$ by Lemma 6.

(III-b-2) If r_2 is ($\to$L) then it must have the following form.

$$\dfrac{B \to C, X_0 \Rightarrow Y, D \qquad B \to C, E, X_0 \Rightarrow Y}{B \to C, D \to E, X_0 \Rightarrow Y}\ (\to\mathrm{L})$$

where $X_0, D \to E = X$. In that case, note that the provable sequent $X \Rightarrow Y, B \to C$ is of the form $X_0, D \to E \Rightarrow Y, B \to C$. We can use Lemma 3 (i) on the proof of the latter to get proofs of both $X_0 \Rightarrow Y, D, B \to C$ and $X_0, E \Rightarrow Y, B \to C$. Thus proceed as follows.

$$\dfrac{\dfrac{X_0 \Rightarrow Y, D, B \to C \qquad B \to C, X_0 \Rightarrow Y, D}{X_0 \Rightarrow Y, D}\ \mathrm{SIH} \qquad \dfrac{X_0, E \Rightarrow Y, B \to C \qquad B \to C, E, X_0 \Rightarrow Y}{X_0, E \Rightarrow Y}\ \mathrm{SIH}}{X_0, D \to E \Rightarrow Y}\ (\to\mathrm{L})$$

Note that both uses of SIH are justified here as the last rule in this proof is effectively an instance of ($\to$L) in PSGLS, hence $\mathrm{mhd}(X_0 \Rightarrow Y, D) < \mathrm{mhd}(X_0, D \to E \Rightarrow Y)$ and $\mathrm{mhd}(X_0, E \Rightarrow Y) < \mathrm{mhd}(X_0, D \to E \Rightarrow Y)$ by Lemma 6.

(III-b-3) If r_2 is (GLR) then it must have the following form.

$$\dfrac{\boxtimes X_0, \Box D \Rightarrow D}{W, B \to C, \Box X_0 \Rightarrow \Box D, \Box Y_0, Z}\ (\mathrm{GLR})$$

where $W, \Box X_0 = X$ and $\Box D, \Box Y_0, Z = Y$. In that case, note that the sequent $X \Rightarrow Y$ is of the form $W, \Box X_0 \Rightarrow \Box D, \Box Y_0, Z$. To obtain a proof of the latter, we apply the rule (GLR) on the premise of r_2 without weakening $B \to C$:

$$\dfrac{\boxtimes X_0, \Box D \Rightarrow D}{W, \Box X_0 \Rightarrow \Box D, \Box Y_0, Z}\ (\mathrm{GLR})$$

(IV) $\mathbf{r_1} =(\to \mathbf{L})$: Then r_1 must have the following form.

$$\frac{X_0 \Rightarrow Y, B, A \qquad C, X_0 \Rightarrow Y, A}{B \to C, X_0 \Rightarrow Y, A} \ (\to\text{L})$$

where $B \to C, X_0 = X$. Thus, we have that the sequents $X \Rightarrow Y$ and $A, X \Rightarrow Y$ are respectively of the form $B \to C, X_0 \Rightarrow Y$ and $A, B \to C, X_0 \Rightarrow Y$. It thus suffices to apply Lemma 3 (i) on the proof of the latter to obtain proofs of both $A, X_0 \Rightarrow Y, B$ and $A, C, X_0 \Rightarrow Y$, and then proceed as follows.

$$\frac{\dfrac{X_0 \Rightarrow Y, B, A \qquad A, X_0 \Rightarrow Y, B}{X_0 \Rightarrow Y, B}\ \text{SIH} \qquad \dfrac{C, X_0 \Rightarrow Y, A \qquad A, C, X_0 \Rightarrow Y}{C, X_0 \Rightarrow Y}\ \text{SIH}}{B \to C, X_0 \Rightarrow Y} \ (\to\text{L})$$

Note that both uses of SIH are justified here as the last rule in this proof is effectively an instance of ($\to$L) in PSGLS, hence $\text{mhd}(X_0 \Rightarrow Y, B) < \text{mhd}(B \to C, X_0 \Rightarrow Y)$ and $\text{mhd}(C, X_0 \Rightarrow Y) < \text{mhd}(B \to C, X_0 \Rightarrow Y)$ by Lemma 6.

(V) $\mathbf{r_1}$ =(GLR): Then we distinguish two cases.

(V-a) A is the diagonal formula in r_1:

$$\frac{\boxtimes X_0, \Box B \Rightarrow B}{W, \Box X_0 \Rightarrow \Box B, \Box Y_0, Z} \ (\text{GLR})$$

where $A = \Box B$ and $W, \Box X_0 = X$ and $\Box Y_0, Z = Y$. Thus, we have that the sequents $X \Rightarrow Y$ and $A, X \Rightarrow Y$ are respectively of the form $W, \Box X_0 \Rightarrow \Box Y_0, Z$ and $\Box B, W, \Box X_0 \Rightarrow \Box Y_0, Z$. We now consider r_2. If r_2 is one of (IdP), ($\bot$L), ($\to$R) and ($\to$L) then respectively proceed as in (I), (II), (III) and (IV) when the cut-formula is not principal in the rules considered by using SIH. We are consequently left to consider the case when r_2 is (GLR). Then r_2 is of the following form:

$$\frac{B, \Box B, \boxtimes X_0, \Box C \Rightarrow C}{W, \Box B, \Box X_0 \Rightarrow \Box C, \Box Y_1, Z} \ (\text{GLR})$$

where $\Box C, \Box Y_1 = \Box Y_0$. In this situation, we distinguish two sub-cases.

(V-a-1) One of the rules (IdP), ($\bot$L) or (IdB) is applicable to $W, \Box X_0 \Rightarrow \Box C, \Box Y_1, Z$, then we are done for the two first cases as it suffices to apply the corresponding rules to obtain a proof of the conclusion of the cut-rule. For the case of (IdB) it suffices to apply Lemma 1.

(V-a-2) None of these rules is applicable to $W, \Box X_0 \Rightarrow \Box C, \Box Y_1, Z$ (NoInit). Then, proceed as follows.

$$\frac{\dfrac{\dfrac{\dfrac{\boxtimes X_0, \Box B \Rightarrow B}{\Box X_0 \Rightarrow \Box B}\ (\text{GLR})}{\boxtimes X_0, \Box C \Rightarrow C, \Box B}\ \text{Lem.2} \qquad \dfrac{\dfrac{\boxtimes X_0, \Box B \Rightarrow B}{\boxtimes X_0, \Box B, \Box C \Rightarrow C, B}\ \text{Lem.2} \qquad B, \Box B, \boxtimes X_0, \Box C \Rightarrow C}{\boxtimes X_0, \Box B, \Box C \Rightarrow C}\ \text{PIH}}{\boxtimes X_0, \Box C \Rightarrow C}\ \text{SIH}}{W, \Box X_0 \Rightarrow \Box C, \Box Y_1, Z} \ (\text{GLR})$$

Note that the use of SIH is justified here as the assumption NoInit ensures that the last rule in this proof is effectively an instance of (GLR) in PSGLS, hence $\text{mhd}(\boxtimes X_0, \Box C \Rightarrow C) < \text{mhd}(W, \Box X_0 \Rightarrow \Box C, \Box Y_1, Z)$ by Lemma 6.

(V-b) A is not the diagonal formula in r_1:

$$\frac{\boxtimes X_0, \Box C \Rightarrow C}{W, \Box X_0 \Rightarrow \Box C, A, \Box Y_0, Z}\ (\mathrm{GLR})$$

where $W, \Box X_0 = X$ and $\Box C, \Box Y_0, Z = Y$. In that case, note that the sequent $X \Rightarrow Y$ is of the form $W, \Box X_0 \Rightarrow \Box C, \Box Y_0, Z$. To obtain a proof of the latter, we apply the rule (GLR) on the premise of r_1 without weakening $\Box B$:

$$\frac{\boxtimes X_0, \Box C \Rightarrow C}{W, \Box X_0 \Rightarrow \Box C, \Box Y_0, Z}\ (\mathrm{GLR})$$

Q.E.D.

The proof of cut-admissibility given here establishes that any topmost cut in a proof in GLS + (cut) is eliminable. By iterating this argument we obtain also cut-elimination for GLS + (cut).

7 Conclusion

We have seen how the termination of backward proof-search can be exploited to obtain cut-elimination. The proof technique used in this paper was first described by Brighton. It is particularly interesting because the termination of backward proof-search is close to a semantic proof of completeness, and the latter is typically much simpler to achieve than a proof of cut-elimination. This makes it particularly interesting to investigate the applicability of this technique to other logics such as Go or intuitionistic GL (using the Dyckhoff calculus for intuitionistic logic since it has terminating backward proof-search).

Our work may appear to beg the following question: if we first need to show semantic cut-free completeness to use this technique, then we already know that every instance of cut is admissible, so, what is the point? Note that this misses the mark. We chose to introduce PSGLS in order to clarify the role of terminating proof-search in the argument, and to demonstrate that the additional notion of regress tree was not essential. In particular, we did not have to show that PSGLS was complete for our purposes.

However, note that it is possible to establish cut-elimination directly without relying on an auxiliary proof calculus such as PSGLS. By isolating the subset of GLS derivations that are also PSGLS derivations, one can use the maximum height on that subset to define the induction measure, and adapt the proofs accordingly.

Acknowledgements. Work supported by the FWF project P33548, CogniGron research center, and the Ubbo Emmius Funds (University of Groningen). Work supported by the FWF projects I 2982 and P 33548.

References

1. Boolos, G.: The Unprovability of Consistency: An Essay in Modal Logic. Cambridge University Press, Cambridge (1979)
2. Borga, M.: On some proof theoretical properties of the modal logic GL. Stud. Logica. **42**(4), 453–459 (1983)
3. Brighton, J.: Cut elimination for GLS using the terminability of its regress process. J. Philos. Log. **45**(2), 147–153 (2016)
4. Dawson, J.E., Goré, R.: Generic methods for formalising sequent calculi applied to provability logic. In: Fermüller, C.G., Voronkov, A. (eds.) LPAR 2010. LNCS, vol. 6397, pp. 263–277. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-16242-8_19
5. Gentzen, G.: Untersuchungen über das logische schließen. II. Math. Zeitschrift **39**, 176–210, 405–431 (1935)
6. Gentzen, G.: Investigations into logical deduction. In: Szabo, M.E. (ed.) The Collected Papers of Gerhard Gentzen, volume 55 of Studies in Logic and the Foundations of Mathematics, pp. 68–131. Elsevier (1969)
7. Goré, R., Ramanayake, R.: Valentini's cut-elimination for provability logic resolved. Rev. Symb. Log. **5**(2), 212–238 (2012)
8. Goré, R., Ramanayake, R.: Cut-elimination for weak Grzegorczyk logic Go. Stud. Log. **102**(1), 1–27 (2014)
9. Indrzejczak, A.: Tautology elimination, cut elimination, and S5. Logic Log. Philos. **26**(4), 461–471 (2017)
10. Mints, G.: Cut elimination for provability logic. In: Collegium Logicum 2005: Cut-Elimination (2005)
11. Negri, S.: Proof analysis in modal logic. J. Philos. Log. **34**(5–6), 507–544 (2005)
12. Negri, S.: Proofs and countermodels in non-classical logics. Log. Univers. **8**(1), 25–60 (2014)
13. Sambin, G., Valentini, S.: The modal logic of provability: the sequential approach. J. Philos. Log. **11**, 311–342 (1982)
14. Sasaki, K.: Löb's axiom and cut-elimination theorem. Acad. Math. Sci. Inf. Eng. Nanzan Univ. **1**, 91–98 (2001)
15. Savateev, Y., Shamkanov, D.: Cut elimination for the weak modal Grzegorczyk logic via non-well-founded proofs. In: Iemhoff, R., Moortgat, M., de Queiroz, R. (eds.) WoLLIC 2019. LNCS, vol. 11541, pp. 569–583. Springer, Heidelberg (2019). https://doi.org/10.1007/978-3-662-59533-6_34
16. Solovay, R.: Provability interpretations of modal logic. Israel J. Math. **25**, 287–304 (1976)
17. Valentini, S.: The modal logic of provability: cut-elimination. J. Philos. Log. **12**, 471–476 (1983)