SIMPLIFIED CUT-ELIMINATION FOR MODAL PROVABILITY LOGICS

ABSTRACT. We give simple, uniform and modular proofs of cut-elimination for GL, Go and Grz. The existing proofs in the literature are highly intricate and specialised.

1. PROBLEM?

What does the cut-elimination procedure produce for the following cut:

$$\frac{p, \Box q \Rightarrow \Box\neg\neg q \qquad \Box\neg\neg q \Rightarrow \Box q, p}{p, \Box q \Rightarrow \Box q, p}$$

It appears to produce this:

$$\frac{q, \Box q \Rightarrow q}{p, \Box q \Rightarrow \Box q, p}$$

But this is NOT a derivation since it breaks the (*) condition?

It may be that you need to keep an eye out for instances of (id) at every step of the proof ? For example, the base case is "if $U, X \Rightarrow Y, V$ is an instance of (id) then stop".

Another possible problem is that mh is not respected upwards by cut!

2. INTRODUCTION

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 [?]. There are several variants with characteristic axioms named after Gödel, Löb and Grzegorczyk:

Name of Logic	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.

Following Gentzen [?], the literature abounds with proofs of cut-elimination for various logics using size of the cut formula and height of the premise derivations. But these measures proved inadequate for proving cut-elimination for GL, so Valentini introduced a third novel measure called width, and showed that cut-elimination for GL now required a triple induction over size, height and width [?].

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 [?, ?, ?]. The question was resolved in Valentini's favour, with all proofs verfied using an interactive theorem prover Isabelle/HOL [?].

His??? proofs for the logic Go are even more intricate, involving ....

Rev, can you please fill in this part as I can't remember the details any more.

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

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

$$(\vee \text{l})\ \frac{A, X \Rightarrow Y \qquad B, X \Rightarrow Y}{A \vee B, X \Rightarrow Y} \qquad (\vee \text{r})\ \frac{X \Rightarrow Y, A, B}{X \Rightarrow Y, A \vee B}$$

$$(\wedge \text{l})\ \frac{A, B, X \Rightarrow Y}{A \wedge B, X \Rightarrow Y} \qquad (\wedge \text{r})\ \frac{X \Rightarrow A, Y \qquad X \Rightarrow B, Y}{X \Rightarrow A \wedge B, Y}$$

$$(\neg \text{l})\ \frac{X \Rightarrow Y, A}{\neg A, X \Rightarrow Y} \qquad (\neg \text{r})\ \frac{A, X \Rightarrow Y}{X \Rightarrow Y, \neg A}$$

$$(\text{GLR})\ \frac{\Box X, X, \Box B \Rightarrow B}{U, \Box X \Rightarrow \Box B, V}\ U \text{ non-boxed multiset}$$

$$(\text{GoR})\ \frac{\Box X, X, \Box(B \to \Box B) \Rightarrow B}{U, \Box X \Rightarrow \Box B, V}\ U \text{ non-boxed multiset}$$

$$(\text{T})\ \frac{A, \Box A, X \Rightarrow Y}{\Box A, X \Rightarrow Y} \text{ need some side-conditions for mh-argument}$$

FIGURE 1. Sequent Rules

Thus the proof-theory of provability logics was significantly more complicated than the proof-theory of numerous other modal logics.

Recently, Brighton [] gave yet another method for proving cut-elimination for GL which is significantly shorter and simpler than any of the existing proofs of cut-elimination in the literature because it uses only a double induction on weight and "maximum height of regress trees".

Here, we finesse Brighton's argument by porting it to a sequent calculus, rather than "regress-trees", and fix some errors in Valentini's original proof. Moreover we show how this argument extends to Go where the only extant proof [?] is highly intricate.

more story and citations

3. Syntax and Axiomatisations of Provability Logics

Modal formulae are constructed from an infinite set $\mathcal{P}$ of atomic formulae using the Backus-Naur form below where $p \in \mathcal{P}$:

$$A, B ::= p \mid \neg A \mid A \to B \mid A \vee B \mid A \wedge B \mid \Box A$$

The $\Diamond$ connective does not appear explicitly but is defined via $\Diamond A := \neg\Box\neg A$.

A *boxed formula* has $\Box$ as its main connective while a *non-boxed* formula does not. A boxed/non-boxed multiset contains only boxed/non-boxed formulae.

The Hilbert-calculus for the basic normal modal logic K is obtained by extending a Hilbert-calulus for classical propositional logic with the axiom $\Box(p \to q) \to (\Box p \to \Box q)$ and the inference rule of neccessitation: from A infer $\Box A$. The various provability logics are then obtained by adding further axioms as shown below:

$$\frac{X, X, Y \Rightarrow Z}{X, Y \Rightarrow Z}\,(\text{lc}) \qquad \frac{X \Rightarrow Y, Z, Z}{X \Rightarrow Y, Z}\,(\text{rc})$$

$$\frac{X \Rightarrow Y}{U, X \Rightarrow Y, V}\,(\text{w}) \quad \frac{X \Rightarrow Y, A \qquad A, U \Rightarrow V}{X, U \Rightarrow Y, V}\,(\text{cut})$$

FIGURE 2. Structural Rules

Name of Logic	Axioms	
GL	$\Box(\Box p \to p) \to \Box p$	
$\mathsf{GL_{lin}}$	$\Box(\Box p \to p) \to \Box p$	$\Box(\Box p \to q) \lor \Box(\Box q \to q)$
Go	$\Box(\Box(p \to \Box p) \to p) \to \Box p$	
Grz	$\Box(\Box(p \to \Box p) \to p) \to \Box p$	$\Box p \to p$

The axiomatisation of Grz given above is indirect since the characteristic axiom $\Box(\Box(p \to \Box p) \to p) \to p$ is actually obtained by chaining the two axioms shown in the table above. This axiomatisation justifies the sequent rules for Grz to follow.

4. Sequent Caluli, Semi-derivations and Maximum Height

A sequent is a tuple (X, Y) of formula multisets, written $X \Rightarrow Y$. Let $X \cap Y$, $X \cup Y$ and $\emptyset$ denote, respectively, the multiset intersection, union and empty multiset. So $\{A\} \cup \{A\}$ is the multiset consisting of two occurrences of A. In sequents, we also write A, X or X, A to stand for $A \cup X$, and similarly write X, Y for $X \cup Y$.

The various sequent calculi are shown below using rules from Figure 1:

Name	Sequent Calculus Rules
PC	(id) ($\neg$L) ($\neg$R) ($\to$L) ($\to$R) ($\land$L) ($\land$R) ($\lor$L) ($\lor$R)
GLS	PC + (GLR)
GoS	PC + (GoR)
GrzS	PC + (GoR) + (T)

Definition 1 (semi-derivation, derivation)**.** *A* semi-derivation *in the sequent calculus is defined inductively as a sequent, or a rule applied upwards to a leaf of a semi-derivation, that additionally satisfies*

($*$) an initial sequent occurs only as a leaf

If every leaf of the semi-derivation is an initial sequent, then it is called a derivation *and its end-sequent (root) is said to be* derivable*.*

Note that ($*$) ensures that only minimal derivations are permitted: no rule is applied upwards to an initial sequent as captured by the (id) rule. Apart from this, the definition of derivation above is the standard one. The following definition of height is also standard.

Definition 2 (height)**.** *The* height *of a semi-derivation is the number of sequents along its longest branch.*

In both (GLR) and (GoR), the purpose of restricting U to a non-boxed multiset is to ensure that all boxed formulae in the antecedent persist upwards in the semi-derivation via $\Box X$. This will be exploited in the proof of Lemma 2 to show that every sequent has a semi-derivation of maximum height.

A single premise sequent rule is *admissible* in a calculus if the conclusion is derivable whenever the premise is derivable. An admissible rule is *height preserving* if the height of the derivation of the conclusion is no greater than that of the derivation of the premise.

It is easy to see via a standard proof theoretic induction on height that the rules (lc), (rc) of contraction and (w) of weakening from Figure 2 are admissible in both GLS and GoS. Using these rules, it is easy to establish the following.

Theorem 1. *The sequent calculi* GLS+(*cut*) *and* GoS+(*cut*) *are sound and complete for the axiomatically formulated logics* GL *and* Go *respectively.*

In the sequel we show how to replace GLS + (cut) and GoS + (cut) with GLS and GoS, respectively, by showing how to transform a derivation with cuts into a derivation of the same end-sequent without cuts.

While height-preserving contraction suffices for most proof theoretic arguments, the use of "maximum height of a semi-derivation of a sequent" as an induction parameter requires a more precise result. A rule is *perfectly height preserving (php) admissible* if the derivations of the conclusion and premise have identical heights.

Lemma 1. *The contraction rule is php-admissible in both* GLS *and* GoS*.*

Will you not need an inversion lemma for this ?

Proof. Induction on the height of a derivation. ⊣

The *complexity* of a sequent s is the number $|s|$ of logical connectives in s.

Lemma 2. *[semi-derivation maximum height mh] In $\mathcal{C}$ ($\mathcal{C} \in \{\mathsf{GLS}, \mathsf{GoS}\}$), every sequent s has a semi-derivation of maximum height, denoted by $mh(s)$. Moreover, for every rule instance with a premise s_1 and conclusion s_0: $mh(s_1) < mh(s_0)$.*

Proof. The proofs for GLS and GoS are almost identical so we only consider GL.

Certainly for every sequent s, there is a semi-derivation ending with s: consider the semi-derivation comprising of the single sequent s. To prove the first claim, we will show that every semi-derivation of s has height bounded by a function in $|s|$.

First note that any boxed formula in the antecedent of a sequent persists in all sequents above it since no rule removes a formula $\Box B$ upwards from the antecedent. This means that the number of (GLR) rules along every branch is bounded by the number of boxed formulae in s, for otherwise, some branch would contain two (GLR) rules with the same diagonal formula $\Box B$ and hence the conclusion of the upper rule would have the form $U, \Box B \Rightarrow \Box B, V$ violating $(*)$.

Observe that the complexity of the premise of a (GLR) rule is at most twice the complexity of the conclusion: the doubling is because the premise contains $\Box X, X, \Box B, B$ whenever the conclusion has the form $\Box X \Rightarrow \Box B$. N.B. In the case of GoS, "at most twice" should be replaced by "at most four times" to account for the larger formula $\Box(B \to \Box B)$ in the premise.

Next, note that the number of propositional rules that can be applied upwards from a given sequent is finite since each such rule results in a sequent with strictly less complexity. We can bound the number of propositional rules that can occur above a sequent s' by its complexity $|s'|$.

Putting this all together, and bounding the number of boxed formulae in s by $|s|$, the height of a semi-derivation of a sequent s in GLS (GoS) is bounded by below left (resp. below right).

$$|s| \cdot 2|s| \cdot 4|s| \cdot 8|s| \cdots 2^{|s|} \cdot |s| \qquad\qquad |s| \cdot 4|s| \cdot 16|s| \cdot 64|s| \cdots 4^{|s|} \cdot |s|$$

For the second claim, suppose s_1/s_0 is an instance of some rule (r) and let $d_{\max}$ be a semi-derivation witnessing the maximum height of the premise s_1. Then the semi-derivation obtained from $d_{\max}$ by applying (r) to its end-sequent is a semi-derivation of s_0 with greater height. ⊣

5. Cut-elimination for GLS

Theorem 2 (Cut-elimination for GLS)**.** *The cut rule is eliminable from* GLS*.*

Proof. Without loss of generality, it suffices to eliminate cut from a derivation concluding with the cut-rule with cut-free left premise $X \Rightarrow Y, A$ and cut-free right premise $A, U \Rightarrow V$. We proceed via primary induction (PI) on the size of the cut-formula and secondary induction (SI) on the maximum height of the combined context $\mathrm{mh}(X, U \Rightarrow Y, V)$. Now, consider the last rule in each premise.

If $X \Rightarrow Y, A$ is an instance of (id) because $X \cap Y \neq \emptyset$, or $A, U \Rightarrow V$ is an instance of (id) because $U \cap V \neq \emptyset$, then so is $X, U \Rightarrow Y, V$.

If $X \Rightarrow Y, A$ is an instance of (id) because $A \in X$ and $A, U \Rightarrow V$ is an instance of (id) because $A \in V$, then so is $X, U \Rightarrow Y, V$.

If the cut-formula A is not principal in either premises of the the last rule, then the cut is shifted upwards following the usual Gentzen reductions. Of course, we need to verify that this does not increase the maximum height. To see that it does not, we illustrate with a generic binary rule (r). Suppose that the left premise of cut ends as below left, so A is not principal in this instance of (r):

$$\dfrac{X' \Rightarrow Y', A \qquad X'' \Rightarrow Y'', A}{X \Rightarrow Y, A}\,(\mathrm{r}) \qquad \dfrac{X', U \Rightarrow Y', V \qquad X'', U \Rightarrow Y'', V}{X, U \Rightarrow Y, V}\,(\mathrm{r})$$

Consider the rule instance of (r) above right, which is what every one of Gentzen's transformations will produce, with the proviso that the derivations of the two premises may now contain an instance of cut on A. Since the cut-formulae is still A, the size of the cut-formula has not changed. Moreover, from part two of Lemma 2, it follows that $\mathrm{mh}(X', U \Rightarrow Y', V) < \mathrm{mh}(X, U \Rightarrow Y, V)$ and $\mathrm{mh}(X'', U \Rightarrow Y'', V) < \mathrm{mh}(X, U \Rightarrow Y, V)$, which means that we can apply the SIH to the premises to obtain cut-free derivations of $X', U \Rightarrow Y', V$ and $X'', U \Rightarrow Y'', V$. Apply (r) to these to obtain a cut-free derivation of $X, U \Rightarrow Y, V$, as shown above right.

If the cut-formula A is principal in both premises by propositional rules then the usual Gentzen reductions suffice, since the new cuts are on smaller formulae and hence the primary induction hypothesis applies. For example the derivation below containing a cut on $A = B \wedge C$

$$\dfrac{\dfrac{X \Rightarrow Y, B \qquad X \Rightarrow Y, C}{X \Rightarrow Y, B \wedge C} \qquad \dfrac{B, C, U \Rightarrow V}{B \wedge C, U \Rightarrow V}}{X, U \Rightarrow Y, V}$$

can be transformed to the following containing a cut on the smaller formula B:

$$\dfrac{\dfrac{X \Rightarrow Y, B \qquad \dfrac{X \Rightarrow Y, C \qquad B, C, U \Rightarrow V}{B, X, U \Rightarrow Y, V}\,(\mathrm{cut})}{X, X, U \Rightarrow Y, Y, V}\,(\mathrm{cut})}{X, U \Rightarrow Y, V}\ \text{Contraction Lemma 1}$$

which falls under the PIH, giving a cut-free derivation of $X, U \Rightarrow Y, V$.

Finally consider when the cut-formula $A = \Box B$ is principal in both premises:

$$\dfrac{\dfrac{\Box X, X, \Box B \Rightarrow B}{L, \Box X \Rightarrow \Box B, M}\,(\text{GLR}) \qquad \dfrac{\Box B, B, \Box U, U, \Box C \Rightarrow C}{R, \Box B, \Box U \Rightarrow \Box C, S}\,(\text{GLR})}{L, R, \Box X, \Box U \Rightarrow \Box C, M, S}\,(\text{cut})$$

The two (GLR) rule premises give an instance of cut with a smaller cut-formula B. The PIH gives a cut-free derivation of $\Box X, X, \Box B, \Box B, \Box U, U, \Box C \Rightarrow C$, and hence by contraction admissibility, a cut-free derivation of $\Box X, X, \Box B, \Box U, U, \Box C \Rightarrow C$.

But observe that the left premise can also be derived without weakening in L and M as shown below left, and that php-contraction on $\Box X, X, \Box X, \Box U, U, \Box C \Rightarrow C$ gives $L, R, \Box X, \Box U \Rightarrow \Box C, M, S$ via (GLR) as shown below right:

$$\dfrac{\Box X, X, \Box B \Rightarrow B}{\Box X \Rightarrow \Box B}\,(\text{GLR}) \qquad \dfrac{\dfrac{\Box X, X, \Box X, \Box U, U, \Box C \Rightarrow C}{\Box X, X, \Box U, U, \Box C \Rightarrow C}\,\text{php-ctr (Lem 1)}}{L, R, \Box X, \Box U \Rightarrow \Box C, M, S}\,(\text{GLR})$$

So $\text{mh}(\Box X, X, \Box X, \Box U, U, \Box C \Rightarrow C) < \text{mh}(L, R, \Box X, \Box U \Rightarrow \Box C, M, S)$. Applying SIH to $\Box X, X, \Box B, \Box U, U, \Box C \Rightarrow C$ and $\Box X \Rightarrow \Box B$ we get a cut-free derivation of $\Box X, X, \Box X, \Box U, U, \Box C \Rightarrow C$. By contraction admissibility there is a cut-free derivation of $\Box X, X, \Box U, U, \Box C \Rightarrow C$. Applying (GLR) we get a cut-free derivation of $L, R, \Box X, \Box U \Rightarrow \Box C, M, S$. ⊣

6. Cut elimination for GoS

If we try to apply directly to GoS the cut-elimination argument for GLS, we encounter the following obstacle in the case when $\Box$ is principal in both premises and we apply the PIH using the premises of each (GoS) rule:

$$\dfrac{\Box X, X, \Box(B \to \Box B) \Rightarrow B \qquad \Box B, B, \Box U, U, \Box(C \to \Box C) \Rightarrow C}{\Box X, X, \Box(B \to \Box B), \Box B, \Box U, U, \Box(C \to \Box C) \Rightarrow C}\,\text{PIH}$$

The issue is the occurrence of $\Box(B \to \Box B)$ in the antecedent that we did not encounter before. To resolve this, we will need a standard invertibility lemma.

I don't see the issue?

Lemma 3. *If $X, A \to B \Rightarrow Y$ is derivable in GoS, then so is $X, B \Rightarrow Y$.*

Proof. Induction on the height of the derivation of $X, A \to B \Rightarrow Y$. ⊣

The following lemma shows that the occurrence of $\Box(B \to \Box B)$ can be removed.

This effectively says that GoR reduces to GLR! That is, every GoS derivation can be transformed into a GLS derivation. Now apply the previous cut-elimination result and then undo the GLS rules into GoR rules!

Lemma 4. *If $X, \Box(C \to \Box C) \Rightarrow Y$ is derivable in GoS, then so is $X, \Box C \Rightarrow Y$.*

Induction on the height of the given derivation.

Base case: If $X \cap Y \neq \emptyset$ then $X, \Box C \Rightarrow Y$ is the required derivation. Otherwise, $Y = Y' \cup \Box(C \to \Box C)$, so use the following derivation of $X, \Box C \Rightarrow Y$, where $X = \Box X_1 \cup U$ with U non-boxed:

$$\dfrac{\dfrac{\dfrac{\Box C \Rightarrow \Box C}{C, \Box X_1, \Box C, C, \Box((C \to \Box C) \to \Box(C \to \Box C)) \Rightarrow \Box C}\,(\text{w})}{\Box X_1, X_1, \Box C, C, \Box((C \to \Box C) \to \Box(C \to \Box C)) \Rightarrow C \to \Box C}\,(\to\text{R})}{\Box X_1, U, \Box C \Rightarrow \Box(C \to \Box C), Y'}\,(\text{GoR})$$

Suppose that the last rule is (GoR) with $X = U \cup \Box X_1$ and $Y = \{\Box B\} \cup V$:

$$\dfrac{\Box X_1, X_1, \Box(C \to \Box C), C \to \Box C, \Box(B \to \Box B) \Rightarrow B}{U, \Box(C \to \Box C), \Box X_1 \Rightarrow \Box B, V}\,(\text{GoR})$$

By IH on the premise, we obtain $\Box X_1, X_1, \Box C, C \to \Box C, \Box(B \to \Box B) \Rightarrow B$. Furthermore, by invertibility of ($\to$l) we get $\Box X_1, X_1, \Box C, \Box C, \Box(B \to \Box B) \Rightarrow B$. By contraction and weakening admissibility we get $\Box X_1, X_1, \Box C, C, \Box(B \to \Box B) \Rightarrow B$. Now apply (GoR) to get $U, \Box X_1, \Box C \Rightarrow \Box B, V$, which is $X, \Box C \Rightarrow Y$.

Suppose the last rule is some other generic (binary) rule below.

$$\frac{X', \Box(C \to \Box C) \Rightarrow Y' \qquad X'', \Box(C \to \Box C) \Rightarrow Y''}{X, \Box(C \to \Box C) \Rightarrow Y}\ (\mathrm{r})$$

Then proceed as follows

$$\cfrac{\cfrac{X', \Box(C \to \Box C) \Rightarrow Y'}{X', \Box C \Rightarrow Y'}\ \mathrm{IH} \qquad \cfrac{X'', \Box(C \to \Box C) \Rightarrow Y''}{X'', \Box C \Rightarrow Y''}\ \mathrm{IH}}{X, \Box C \Rightarrow Y}\ (\mathrm{r})$$

Theorem 3 (Cut-elimination for GoS)**.** *The cut rule is eliminable from* GoS.

Proof. Once again, without loss of generality, it suffices to eliminate cut from a derivation concluding with the cut-rule with cut-free left premise $X \Rightarrow Y, A$ and cut-free right premise $A, U \Rightarrow V$. Primary induction on the size of the cut-formula and secondary induction on the maximum height of the combined contexts $\mathrm{mh}(X, U \Rightarrow Y, V)$. The proof is identical to that of Theorem 2 except the following case.

Consider the case when the cut formula $\Box B$ is principal in both premises.

$$\cfrac{\cfrac{\Box X, X, \Box(B \to \Box B) \Rightarrow B}{L, \Box X \Rightarrow \Box B, M}\ (\mathrm{GoR}) \qquad \cfrac{\Box B, B, \Box U, U, \Box(C \to \Box C) \Rightarrow C}{R, \Box B, \Box U \Rightarrow \Box C, S}\ (\mathrm{GoR})}{L, R, \Box X, \Box U \Rightarrow \Box C, M, S}$$

By PIH using the premises of each (GoR) rule, we obtain a cut-free derivation of $\Box X, X, \Box(B \to \Box B), \Box B, \Box U, U, \Box(C \to \Box C) \Rightarrow$ C. By Lemma 4 we have a cut-free derivation of $\Box X, X, \Box B, \Box B, \Box U, U, \Box(C \to \Box C) \Rightarrow C$. Hence by contraction admissibility, a cut-free derivation of $\Box X, X, \Box B, \Box U, U, \Box(C \to \Box C) \Rightarrow C$.

Observe that

$$\cfrac{\cfrac{\Box X, X, \Box X, \Box U, U, \Box(C \to \Box C) \Rightarrow C}{\Box X, X, \Box U, U, \Box(C \to \Box C) \Rightarrow C}\ \text{php contraction (Lem. 1)}}{L, R\Box X, \Box U \Rightarrow \Box C, M, S}\ (\mathrm{GoR})$$

So $\mathrm{mh}(\Box X, X, \Box X, \Box U, U, \Box(C \to \Box C) \Rightarrow C) < \mathrm{mh}(L, R, \Box X, \Box U \Rightarrow \Box C, M, S)$. Applying SIH to $\Box X, X, \Box B, \Box U, U, \Box(C \to \Box C) \Rightarrow C$ and $\Box X \Rightarrow \Box B$ we get a cut-free derivation of $\Box X, X, \Box X, \Box U, U, \Box(C \to \Box C) \Rightarrow C$. By contraction admissibility there is a cut-free derivation of $\Box X, X, \Box U, U, \Box(C \to \Box C) \Rightarrow C$. Applying (GoR) we get a cut-free derivation of $\Box X, \Box U \Rightarrow \Box C$. ⊣

References

E-mail address: revantha@logic.at

Tuesday 19th February, 2019. 12:04