THE REVIEW OF SYMBOLIC LOGIC
Volume 5, Number 2, June 2012

VALENTINI'S CUT-ELIMINATION FOR PROVABILITY LOGIC RESOLVED

RAJEEV GORÉ
Logic and Computation Group, Research School of Computer Science,
The Australian National University
and
REVANTHA RAMANAYAKE
CNRS LIX, École Polytechnique

Abstract. Valentini (1983) has presented a proof of cut-elimination for provability logic GL for a sequent calculus using sequents built from sets as opposed to multisets, thus avoiding an explicit contraction rule. From a formal point of view, it is more syntactic and satisfying to explicitly identify the applications of the contraction rule that are 'hidden' in proofs of cut-elimination for such sequent calculi. There is often an underlying assumption that the move to a proof of cut-elimination for sequents built from multisets is straightforward. Recently, however, it has been claimed that Valentini's arguments to eliminate cut do not terminate when applied to a multiset formulation of the calculus with an explicit rule of contraction. The claim has led to much confusion and various authors have sought new proofs of cut-elimination for GL in a multiset setting.

Here we refute this claim by placing Valentini's arguments in a formal setting and proving cut-elimination for sequents built from multisets. The use of sequents built from multisets enables us to accurately account for the interplay between the weakening and contraction rules. Furthermore, Valentini's original proof relies on a novel induction parameter called "width" which is computed 'globally'. It is difficult to verify the correctness of his induction argument based on "width." In our formulation however, verification of the induction argument is straightforward. Finally, the multiset setting also introduces a new complication in the case of contractions above cut when the cut-formula is boxed. We deal with this using a new transformation based on Valentini's original arguments.

Finally, we discuss the possibility of adapting this cut-elimination procedure to other logics axiomatizable by formulae of a syntactically similar form to the GL axiom.

§1. Introduction. The sequent calculus was first introduced by Gentzen (1969) as a tool to study the metamathematical properties of his normal deduction systems for first-order intuitionistic and classical logic. Subsequently, Gentzen's sequent calculus and its generalizations such as hypersequent calculi (Pottinger, 1983; Avron, 1996), display calculi (Belnap, 1982), and the calculus of structures (Guglielmi, 2007) have been employed to present a variety of different logics including modal, relevant, and substructural logics.

Gentzen included the cut-rule below in his sequent calculi in order to prove completeness with respect to the corresponding natural deduction systems:

$$\frac{X \Rightarrow Y, A \qquad A, U \Rightarrow W}{X, U \Rightarrow Y, W}\ cut$$

Received: August 25, 2010.

© Association for Symbolic Logic, 2012

doi:10.1017/S1755020311000323

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

The cut-rule simplifies proofs of completeness with respect to Hilbert calculi as well, mimicking the rule of *modus ponens*. Furthermore, the cut-rule can considerably shorten the length of a derivation.

However, in many other ways the presence of the cut-rule is undesirable. Read upwards, from the conclusion to the premises, the cut-rule can introduce an arbitrary formula into the derivation, causing difficulties in proof search. Showing *how* any given derivation containing applications of the cut-rule can be transformed into a derivation of the same sequent without using the cut-rule is known as *cut-elimination*. Cut-elimination yields a host of metamathematical results such as completeness, consistency, decidability, and interpolation (Takeuti, 1987). In the case of the original sequent calculi LJ and LK, the cut-elimination theorem—often called Gentzen's *Hauptsatz* or 'main theorem'—corresponds to the normalization result for the corresponding natural deduction systems. Informally we can think of an application of the cut-rule as corresponding to the use of a lemma in a proof. Then cut-elimination corresponds to the statement that any proof using lemmas can be rewritten effectively to obtain a proof of the same theorem that does not use any lemmas. Under this interpretation, the failure of cut-elimination would indicate that the proof calculus fails this rather reasonable requirement.

Extending Gentzen's *Hauptsatz* by presenting sequent calculi with cut-elimination for other logics is a main challenge in proof theory. It is important to stress here that cut-elimination is not a property of the logic, but rather a property of its proof calculus. For example, by taking all the derivable sequents as axioms it is possible to trivially delete the cut-rule from the sequent calculus. However the aim is to obtain a sequent calculus with nice technical properties. Observe that there is a weaker result called *cut-redundancy* which states that the cut-rule is redundant in a sequent calculus without specifying *how* to eliminate the cuts in a given derivation. Proofs of cut-redundancy usually appeal to the semantics of the logic. However, from a structural proof theory perspective, it is the syntactic constructive procedure relating the original derivation with a cut-free derivation of the identical sequent that is of interest. Clearly, cut-elimination implies cut-redundancy.

In this paper we look at cut-elimination for provability logic GL. The provability logic GL is obtained by adding Löb's axiom $\Box(\Box A \supset A) \supset \Box A$ to the standard Hilbert calculus for propositional normal modal logic K (Sambin & Valentini, 1982). The logic is interesting in its own right, largely due to its connections with provability in Peano arithmetic. Interpreting the modal operator $\Box A$ as the provability predicate "A is provable in Peano arithmetic," it is well-known that GL is complete with respect to the formal provability interpretation in Peano arithmetic (see Solovay, 1976). For an introduction to provability logic see Švejdar (1999).

1.1. History of cut-elimination for GL. The problem of cut-elimination for GL has its own chequered history. Leivant (1981) proposed a proof of cut-elimination for a sequent calculus for GL. Valentini (1983) soon described a counterexample to this proof, proposing a more complicated proof for the sequent calculus GLS_V for GL. The calculus GLS_V is a sequent calculus for classical propositional logic together with a single modal rule GLR. Valentini's proof appears to be the first proof of cut-elimination for a sequent calculus for GL and relies on a complicated transformation justified by a Gentzen-style induction on the degree of the cut-formula and the cut-height, as well as a new induction parameter—the width of a cut-formula. Roughly speaking, the width of a cut-instance is the number of GLR rule instances above that cut which contain a parametric ancestor of the cut-formula in their conclusion. However, Valentini's proof is very brief, informal, and difficult to check. For example, he only considers a cut-instance where the cut-formula is left and right

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

principal by the GLR rule (the Sambin Normal Form), noting that "the presence of the new parameter [width] does not affect the [remainder of the standard cut-elimination proof]" (Valentini, 1983). While it is true that the standard transformations appropriately reduce the degree and/or cut-height, he fails to observe that these transformations can sometimes increase the width of lower cuts, casting doubt on the validity of the induction. A careful study of the proposed transformation is required to confirm that the proof is not affected (see Remark 5.5).

Several other solutions for cut-elimination have been proposed. Borga (1983) presented one solution, while Sasaki (2001) described a proof for a sequent calculus very similar to GLS_V, relying on cut-elimination for $K4$. Note that only Leivant and Valentini used traditional Gentzen-style methods involving an induction over the degree of the cut-formula and the cut-height.

All four authors used sequents $X \Rightarrow Y$ where X and Y are *sets*, so these calculi did not require a rule of contraction as there is no notion of a set containing an element multiple times (unlike a multiset where the number of occurrences is important). Thus the following instance of the $L\wedge$ rule is legal in GLS_V even though it 'hides' a contraction on $P \wedge Q$:

$$\frac{P \wedge Q, P \Rightarrow R}{P \wedge Q \Rightarrow R}\, L\wedge$$

From the syntactic viewpoint adopted in structural proof theory, it is more satisfying to explicitly identify the contractions that are 'hidden' in these set-based proofs of cut-elimination. The appropriate formalization to understand the reliance on contraction is to use sequents built from multisets and add the rules of contraction to the calculus. This enables us to identify and eliminate implicit occurrences of contraction. Using sequents built from multisets the derivation above becomes the following, where LC denotes the 'left contraction' rule:

$$\dfrac{\dfrac{P \wedge Q, P \Rightarrow R}{P \wedge Q, P \wedge Q \Rightarrow R}\, L\wedge}{P \wedge Q \Rightarrow R}\, LC$$

However, the presence of the contraction rules often pose new issues for cut-elimination. Gentzen (1969), in his original proof of the *Hauptsatz* for the classical sequent calculus LK (he used sequents built from lists, so a contraction rule was still required), introduced the *multicut* rule, below, in order to eliminate cut from a derivation in the problematic case where the contraction rule occurs immediately above the cut. In the following, $n, m > 0$:

$$\frac{X \Rightarrow Y, A^n \qquad A^m, U \Rightarrow W}{X, U \Rightarrow Y, W}\, mcut$$

Notice that the multicut rule implicitly executes contraction on the multisets A^n and A^m, thus combining the structural rules of contraction and cut. This is undesirable as it hinders our ability to analyze the independent 'effect' of each rule, and so the use of multicut is not a satisfactory solution for understanding the reliance on the contraction rule. In recognition of this, numerous proofs of cut-elimination for LK for sequents built from multisets and avoiding multicut have appeared in the literature (see, e.g., von Plato, 2001; Borisavljević *et al.*, 2000; Bimbó, 2007) to deal with the case of contractions above cut. In the case of GL, it turns out that additional complications also arise when formulating Valentini's arguments in a multiset setting, for example, due to the interplay between weakening and contraction rules (see Remark 3.13).

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

Thus the move to a proof of cut-elimination for sequents built from multisets is not straightforward. Moen (2001) attempted to lift Valentini's set-based arguments to obtain a proof for sequents built from multisets before concluding that this was not possible. Specifically, he presents a concrete derivation ϵ containing cut, and claims that a multiset formulation of Valentini's argument does not terminate when applied to ϵ. Not surprisingly, this claim has ignited the search for new proofs of cut-elimination in a Gentzen-style multiset setting for GL.

In response, Negri (2005) and Mints (2005) proposed two different solutions. Negri uses a nonstandard multiset sequent calculus in which sequents are built from multisets of labeled formulae of the form $x : A$, where A is a traditional formula and x is an explicit name for a Kripke world. She shows that contraction is height-preserving admissible in this calculus and uses this to handle contractions above cut in her cut-elimination argument. In our view, the use of semantic information in the calculus deviates from a purely proof-theoretic approach. Mints (2005) solves the problem using a sequent calculus similar to the multiset formulation of GLS_V, but does not state how to handle contractions above cut.

Poggiolesi (2009) presented a proof of cut-elimination for a *tree-hypersequent* calculus for GL. A tree-hypersequent is built using some number of (Gentzen) sequents that are ordered and nested using new metalevel symbols '/' and ';', and yields a structure that can be read in terms of a tree. In particular, the order of the sequents (as determined by the '/', ';' symbols) is important. Contrast this with traditional hypersequents (see Avron, 1996) built from some number of traditional sequents separated by '/', where the order of the sequents is *not* important. Poggiolesi claims that the tree-hypersequent calculus for GL has all the advantages of Negri's calculus, and in addition does not make use of any semantic elements. However it appears that tree-hypersequents 'hide' the labeling for Kripke worlds by making use of the ordering and nesting induced by '/' and ';'. In other words, the reliance on the ordering of the sequents means that tree-hypersequents contain similarly explicit semantic information, although disguised through the use of less suggestive symbols.

To obtain a satisfactory proof of cut-elimination for GL for sequents built from multisets, there are two issues to consider:

1. formalize "width" more precisely to clarify Valentini's original definition, and check whether it is a suitable induction measure, and
2. determine whether Valentini's arguments can be used to obtain a purely syntactic proof of cut-elimination in a *multiset* setting.

By dealing with these two issues we can show that it is not necessary to move to a new formalism in order to obtain a proof of cut-elimination for GL for sequents built from multisets.

1.2. The result. Our contribution is as follows: we have successfully translated Valentini's set-based arguments for cut-elimination to a sequent calculus built from multisets. To this end, we have formalized the notion of parametric ancestor and width to correspond intuitively with Valentini's original definition. With this formalization we show that Valentini's arguments can be applied in the multiset setting, noting that although certain transformations may increase the width of lower cuts, this does not affect the proof. In the case where the last rule in either premise derivation of the cut-rule is a contraction on the cut-formula, we avoid the multicut rule by using von Plato's (2001) arguments when the cut-formula is not boxed, and a new argument for the case when

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

the cut-formula is boxed. Thus we obtain a purely syntactic proof of cut-elimination in a multiset setting. We also believe that we have identified a mistake in Moen's claim that Valentini's arguments (in a multiset setting) do not terminate. It appears that Moen has not used a faithful representation of Valentini's arguments for the inductive case, but instead a transformation he titles Val-II(core) that is already known to be insufficient (Sambin & Valentini, 1982). We discuss this further in §6. Of course, the incorrectness of Moen's claim does not imply the correctness of Valentini's arguments in a multiset setting. Indeed the whole point is that complications do arise in the multiset setting, and that these have to be dealt with carefully.

We remind the reader that it is trivial to show cut-redundancy for both set and multiset sequent calculus formulations for GL by proving that the calculus without the cut-rule is sound and complete for the Kripke semantics of GL. However, our purpose here was to resolve the claim about the failure of cut-elimination based on Valentini's arguments for a sequent calculus built with multisets.

In §7 we discuss the possibility of adapting the cut-elimination procedure for GLS to other logics whose axiomatizations over the basic normal modal logic K contain an axiom of a syntactically similar form to Löb's axiom.

This paper is a longer version of Goré & Ramanayake (2008).

§2. Preliminaries. We use the letters $P, Q, \ldots$ to denote propositional variables. Formulae are defined in the usual way in terms of propositional variables, the logical constant $\bot$ and the logical connectives $\wedge$ (conjunction), $\vee$ (disjunction), $\supset$ (implication), and $\Box$ (necessity, or in this context, provability). Formulae are denoted by $A, B, \ldots$. Multisets of formulae are denoted by X, Y, U, V, W, G and also as a list of comma-separated formula enclosed in "$\langle$" and "$\rangle$." A formula A is said to be *boxed* if it is of the form $\Box B$ for some formula B and is *not boxed* otherwise. The relation '$\equiv$' is used to denote syntactic equality. Let X be the multiset $A_1, \ldots, A_n$. Then we define the multiset $\Box X$ to be $\Box A_1, \ldots, \Box A_n$. Furthermore $B \in X$ iff $B \equiv A_i$ for some $1 \leq i \leq n$. The negation of $B \in X$ is denoted by $B \notin X$. The notation $(A)^m$ or A^m denotes m comma-separated occurrence of A.

A *sequent* is a tuple (X, Y) of multisets X and Y of formulae and is written $X \Rightarrow Y$. We sometimes use $\mathcal{S}$ or $\mathcal{S}'$ to denote a sequent. The multiset X (Y) is called the antecedent (succedent). The multiset sequent calculus we use here is called GLS (Table 1). For the logical and structural rules in GLS, the multisets X and Y are called the *context*. In the conclusion of each of these rules, the formula occurrence not in the context is called the *principal formula*. This follows standard practice (see Troelstra & Schwichtenberg, 2000). For the GLR rule, each formula in $\Box X, X, \Box B, B$ is called a principal formula. The $\Box B$ in the succedent of the conclusion of the GLR rule is called the *diagonal formula* (Valentini, 1983) and is of course boxed. In the cut-rule, the formula D is the *cut-formula*. A rule with one premise (two premises) is called a unary (binary) rule.

A binary rule where the context in both premises is required to be identical is called an *additive* binary rule (e.g., $L\vee$, $R\wedge$). A binary rule where the context in each premise need not be identical is called a *multiplicative* binary rule (e.g., *cut*). The term context sharing (context independent) is also used to refer to an additive (multiplicative) rule (see Troelstra & Schwichtenberg, 2000).

Note, we have deleted the initial sequent $\bot \Rightarrow \bot$ and the $\bot$-rule that appears in GLS_V. As Švejdar (1999) observes, it is not necessary to include $\bot$ although its presence can be convenient from a semantic viewpoint. As our concerns here are proof theoretic we shall not require it. We have also replaced the multiplicative $L\supset$ in GLS_V with an additive

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

Initial sequents: $A \Rightarrow A$ for each formula A

Logical rules:

$$\frac{X \Rightarrow Y, A}{X, \neg A \Rightarrow Y}\, L\neg \qquad \frac{A, X \Rightarrow Y}{X \Rightarrow Y, \neg A}\, R\neg$$

$$\frac{A_i, X \Rightarrow Y}{A_1 \wedge A_2, X \Rightarrow Y}\, L\wedge \qquad \frac{X \Rightarrow Y, A_1 \qquad X \Rightarrow Y, A_2}{X \Rightarrow Y, A_1 \wedge A_2}\, R\wedge$$

$$\frac{A_1, X \Rightarrow Y \qquad A_2, X \Rightarrow Y}{A_1 \vee A_2, X \Rightarrow Y}\, L\vee \qquad \frac{X \Rightarrow Y, A_i}{X \Rightarrow Y, A_1 \vee A_2}\, R\vee$$

$$\frac{X \Rightarrow Y, A \qquad B, X \Rightarrow Y}{A \supset B, X \Rightarrow Y}\, L\supset \qquad \frac{A, X \Rightarrow Y, B}{X \Rightarrow Y, A \supset B}\, R\supset$$

Modal rule: $$\frac{\Box X, X, \Box B \Rightarrow B}{\Box X \Rightarrow \Box B}\, GLR$$

Structural rules:

$$\frac{X \Rightarrow Y}{A, X \Rightarrow Y}\, LW \qquad \frac{X \Rightarrow Y}{X \Rightarrow Y, A}\, RW$$

$$\frac{A, A, X \Rightarrow Y}{A, X \Rightarrow Y}\, LC \qquad \frac{X \Rightarrow Y, A, A}{X \Rightarrow Y, A}\, RC$$

Cut-rule: $$\frac{X \Rightarrow Y, D \qquad D, U \Rightarrow W}{X, U \Rightarrow Y, W}\, cut$$

Table 1. *The sequent calculus GLS*

version. As all the other logical rules in GLS are additive, it seems appropriate to use an additive $L\supset$. We observe that the definitions and proofs in this paper apply, with slight amendment, to a sequent calculus built from multisets that is obtained directly from GLS_V.

A *derivation* (in GLS) is defined recursively with reference to Table 1 as:

(i) an initial sequent $A \Rightarrow A$ for any formula A is a derivation, and
(ii) an application of a logical, modal, structural, or cut-rule to derivations concluding its premise(s) is a derivation.

This is the standard definition. Viewing a derivation as a tree, we call the root of the tree the *end-sequent* of the derivation. If there is a derivation with end-sequent $X \Rightarrow Y$ we say that $X \Rightarrow Y$ is *derivable* in GLS. Let $\bigwedge X$ ($\bigvee Y$) denote the conjunction (disjunction) of all formula occurrences in X (Y). Interpreting the sequent $X \Rightarrow Y$ as the formula $\bigwedge X \supset \bigvee Y$, from Sambin & Valentini (1982) we see that derivability in GLS is sound and complete wrt GL.

We write $\{\pi\}_1^r/^\rho X \Rightarrow Y$ to denote the derivation

$$\frac{\pi_1 \qquad \ldots \qquad \pi_r}{X \Rightarrow Y}\, \rho$$

where ρ is an instance of a rule with r premises. We refer to $\pi_1, \ldots, \pi_r$ as the *premise derivations* of ρ. If ρ is unary (binary) then $r = 1$ ($r = 2$). Rather than $\{\pi\}_1^1$ and $\{\pi\}_1^2$, we write, respectively, "π_1" and "$\pi_1 \quad \pi_2$."

Let ρ be some rule occurrence in a derivation τ. Then $\rho(A)$ indicates that the principal formula is A and $\rho^*(X)$ denotes some number (≥ 0) of applications of ρ that make

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

each formula occurrence (including multiple formula occurrences) in the multiset X a principal formula. To identify a rule-occurrence in the text we occasionally use subscripts, for example, GLR_1, cut_0.

A derivation τ is *cut-free* if τ contains no instances of the cut-rule. A cut-instance is said to be *topmost* if its premise derivations are cut-free.

DEFINITION 2.1 (n-ary GLR rule) *Given a derivation τ, an instance ρ of the GLR rule appearing in τ is n-ary if there are exactly $n-1$ GLR rule instances on the path between the conclusion sequent of ρ and the end-sequent of τ.*

EXAMPLE 2.2. *Consider the following derivation τ:*

$$\dfrac{\dfrac{\dfrac{\Box(\Box A \wedge \Box B), \Box A \wedge \Box B, \Box A \Rightarrow A}{\Box(\Box A \wedge \Box B) \Rightarrow \Box A}\,\rho_1 \quad \dfrac{\Box(\Box A \wedge \Box B), \Box A \wedge \Box B, \Box B \Rightarrow B}{\Box(\Box A \wedge \Box B) \Rightarrow \Box B}\,\rho_2}{\Box(\Box A \wedge \Box B) \Rightarrow \Box A \wedge \Box B}\,R\wedge}{\Rightarrow \Box(\Box A \wedge \Box B)}\,\rho_3$$

Note that ρ_1, ρ_2, and ρ_3 are instances of the GLR rule. By inspection, ρ_1 is a 2-ary GLR instance because there is a single GLR rule instance (ρ_3) on the path between the conclusion sequent of ρ_1 and the end-sequent $\Rightarrow \Box(\Box A \wedge \Box B)$. Similarly, ρ_2 is a 2-ary GLR rule instance. Finally, because the conclusion sequent of ρ_3 is simultaneously the end-sequent of τ, it follows that ρ_3 is a 1-ary GLR rule instance.

Let $GLR(n, \tau)$ denote the number of n-ary GLR rules in τ. Next we define the height, cut-height, and degree of a formula in the standard manner.

DEFINITION 2.3 (height, cut-height, degree) *The height $s(\tau)$ of a derivation τ is the greatest number of successive applications of rules in it plus one. The cut-height h of an instance of the cut-rule with premise derivations τ_1 and τ_2 is $s(\tau_1) + s(\tau_2)$. The degree $deg(A)$ of a formula A is defined as the number of symbol occurrences in A from $\{\Box, \neg, \wedge, \vee, \supset\}$ plus one.*

§3. Generalizing the notion of derivation. To formalize the notion of width we need a more general structure than a derivation. The structure we have in mind can be obtained from a derivation τ by deleting a proper subderivation τ' in τ. We call this structure a *stub-derivation*. To emphasize the point of deletion we use the annotation stub.

Formally a stub-derivation (in GLS) is defined recursively with reference to Table 1 as follows:

(i) an initial sequent $A \Rightarrow A$ for any formula A is a stub-derivation, and
(ii) for any sequent $\mathcal{S}$ and stub-derivation π, each of

(a) $\dfrac{\mathsf{stub}}{\mathcal{S}}$ (b) $\dfrac{\mathsf{stub} \quad \pi}{\mathcal{S}}$ (c) $\dfrac{\pi \quad \mathsf{stub}}{\mathcal{S}}$

is a stub-derivation, and
(iii) an application of a logical, modal, structural, or cut-rule to stub-derivations concluding its premise(s) is a stub-derivation.

Viewing a stub-derivation τ as a tree, we call the root of the tree the *end-sequent* of the stub-derivation (denoted $ES(\tau)$). The leaves of the tree are called the *top-sequents*. Clearly a derivation is a stub-derivation in which every top-sequent is an initial sequent. Thus a stub-derivation generalizes the notion of a derivation.

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

We use the term 'stub-instance' to refer to an occurrence of either $\mathsf{stub}/\mathcal{S}$ or $\mathsf{stub}\quad \pi/\mathcal{S}$ or $\pi\quad \mathsf{stub}/\mathcal{S}$. An *sstub-derivation* (read: single stub-derivation) is a stub-derivation containing exactly one stub-instance. We write $d[\mathsf{stub}]$ instead of d, to remind the reader that the structure contains exactly one stub-instance.

Let d' be a derivation with end-sequent $\mathcal{S}'$, let $d[\mathsf{stub}]$ be an sstub-derivation with an occurrence of one of the following:

$$\mathsf{stub}/\mathcal{S} \qquad\qquad \mathsf{stub}\quad \pi/\mathcal{S} \qquad\qquad \pi\quad \mathsf{stub}/\mathcal{S}$$

Now suppose that each of the following is respectively a legal instance of some logical or structural rule ρ:

$$\mathcal{S}'/^{\rho}\mathcal{S} \qquad\qquad \mathcal{S}'\quad ES(\pi)/\mathcal{S} \qquad\qquad ES(\pi)\quad \mathcal{S}'/\mathcal{S}$$

We say that $d[\mathsf{stub}]$ and d' are *compatible* and write $d[\mathsf{stub}] \leftarrowtail d'$ to denote the following trees, respectively obtained by "attaching" the tree d' to the tree $d[\mathsf{stub}]$ at the node stub under rule ρ:

$$\frac{d'}{\mathcal{S}}\,\rho \qquad\qquad \frac{d' \qquad \pi}{\mathcal{S}}\,\rho \qquad\qquad \frac{\pi \qquad d'}{\mathcal{S}}\,\rho$$

We refer to ρ as a *binding rule* for $d[\mathsf{stub}]$ and d'.

By permitting formula occurrences in a (stub-)derivation to contain $*$ or $\circ$ decorations, we define an *annotated (stub-)derivation*. The forgetful map $\lfloor\cdot\rfloor$ maps an annotated stub-derivation to the stub-derivation obtained by erasing all $*$ and $\circ$ decorations. Clearly $\lfloor\cdot\rfloor$ maps an annotated derivation to a derivation. A *transformed (stub-)derivation* τ' is a (stub-)derivation that is obtained from some existing (stub-)derivation τ by syntactic transformation. We write $A^{\circ n}$ or A^{*n} to mean n occurrences of the formula A° or A^{*} respectively.

Formally a stub-derivation and an annotated stub-derivation are different structures. Because these structures are very similar, for economy of space we introduce definitions and prove results for stub-derivations alone and note, whenever applicable, that the definitions and results can be extended to annotated stub-derivations.

EXAMPLE 3.1. *Let us denote the sstub-derivation at below left by* $d[\mathsf{stub}]$ *and the derivation at below right by* d'*:*

$$\dfrac{\dfrac{\mathsf{stub}}{B \Rightarrow A \supset B} \qquad A \supset B \Rightarrow A \supset B}{B \vee (A \supset B) \Rightarrow A \supset B}\, L\vee \qquad\qquad \dfrac{B \Rightarrow B}{A, B \Rightarrow B}\, LW$$

Observe that $d[\mathsf{stub}]$ *has a stub-instance of type* $\mathsf{stub}/\mathcal{S}$*, with* $\mathcal{S} \equiv B \Rightarrow A \supset B$*, and* d' *has end-sequent* $\mathcal{S}' \equiv A, B \Rightarrow B$*.*

Because $\mathcal{S}'/\mathcal{S}$ *is an instance of* $R\supset$*, the structures* $d[\mathsf{stub}]$ *and* d' *are compatible. The derivation* $d[\mathsf{stub}] \leftarrowtail d'$ *is as shown below and the binding rule is* $R\supset$*:*

$$\dfrac{\dfrac{\dfrac{B \Rightarrow B}{A, B \Rightarrow B}\, LW}{B \Rightarrow A \supset B}\, R\supset \qquad A \supset B \Rightarrow A \supset B}{B \vee (A \supset B) \Rightarrow A \supset B}\, L\vee$$

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

EXAMPLE 3.2. *Let us denote the sstub-derivation at below left by* $d[\mathsf{stub}]$ *and the derivation at below right by* d'*:*

$$\dfrac{\mathsf{stub} \qquad A \supset B \Rightarrow A \supset B}{B \vee (A \supset B) \Rightarrow A \supset B} \qquad\qquad \dfrac{\dfrac{B \Rightarrow B}{A, B \Rightarrow B}\, LW}{B \Rightarrow A \supset B}\, R\supset$$

Observe that $d[\mathsf{stub}]$ *has a stub-instance of type* $\mathsf{stub} \quad \tau/\mathcal{S}$*, with* $\mathcal{S} \equiv B \vee (A \supset B) \Rightarrow A \supset B$*, and* d' *has endsequent* $\mathcal{S}' \equiv B \Rightarrow A \supset B$*.*

Since $\mathcal{S}' \quad A \supset B \Rightarrow A \supset B/\mathcal{S}$ *is an instance of* $L\vee$*, the structures* $d[\mathsf{stub}]$ *and* d' *are compatible. The derivation* $d[\mathsf{stub}] \leftarrow\!\!\!\shortmid\, d'$ *is identical to that obtained in Example 3.1, although here the binding rule is* $L\vee$*.*

DEFINITION 3.3 *Let* τ *be a stub-derivation and* G *a formula multiset. The antecedent operator* $\oplus$: *stub-derivation* $\times$ *formula multiset* $\mapsto$ *stub-derivation is defined as follows:*

Case $G = \{\}$*: let* $\tau \oplus G = \tau$
Case $G \neq \{\}$*: define* $\tau \oplus G$ *recursively on* τ *as follows*

1. *initial sequent:* $(A \Rightarrow A) \oplus G = (A \Rightarrow A/^{LW^*(G)}A, G \Rightarrow A)$
2. *stub-instance:*
   (a) $(\mathsf{stub}/X \Rightarrow Y) \oplus G = (\mathsf{stub}/X, G \Rightarrow Y)$
   (b) $(\mathsf{stub} \quad \pi/X \Rightarrow Y) \oplus G = (\mathsf{stub} \quad \pi \oplus G/X, G \Rightarrow Y)$
   (c) $(\pi \quad \mathsf{stub}/X \Rightarrow Y) \oplus G = (\pi \oplus G \quad \mathsf{stub}/X, G \Rightarrow Y)$
3. *unary non-GLR rule:* $(\pi/X \Rightarrow Y) \oplus G = (\pi \oplus G/X, G \Rightarrow Y)$
4. *GLR rule:* $(\pi/^{GLR}X \Rightarrow Y) \oplus G = (\pi/^{GLR}X \Rightarrow Y)/^{LW^*(G)}X, G \Rightarrow Y$
5. *binary additive rule:* $(\pi_1 \; \pi_2/X \Rightarrow Y) \oplus G = (\pi_1 \oplus G \; \pi_2 \oplus G/X, G \Rightarrow Y)$
6. *cut-rule:* $(\pi_1 \quad \pi_2/^{cut}X \Rightarrow Y) \oplus G = (\pi_1 \oplus G \quad \pi_2/^{cut}X, G \Rightarrow Y)$.

That $\oplus$ maps into the set of stub-derivations follows by inspection of the definition. Notice that the recursion terminates at an initial sequent, stub-instance, or a GLR rule. The operator $\oplus$ will bind stronger that $\leftarrow\!\!\!\shortmid$.

LEMMA 3.4. *If* d *is a stub-derivation and* G *is a formula multiset, then* $d \oplus G$ *is a stub-derivation. Furthermore, if* d *is in fact an sstub-derivation* $d[\mathsf{stub}]$*, then* $d[\mathsf{stub}] \oplus G$ *is an sstub-derivation.*

Proof. The result follows immediately from Definition 3.3. □

EXAMPLE 3.5. *Refer to the sstub-derivation* $d[\mathsf{stub}]$ *in Example 3.1. If* G *is a nonempty formula multiset, then* $d[\mathsf{stub}] \oplus G$ *is the stub-derivation:*

$$\dfrac{\dfrac{\mathsf{stub}}{B, G \Rightarrow A \supset B} \qquad \dfrac{A \supset B \Rightarrow A \supset B}{A \supset B, G \Rightarrow A \supset B}\, LW^*(G)}{B \vee (A \supset B), G \Rightarrow A \supset B}\, L\vee$$

By observation, we can confirm that $d[\mathsf{stub}] \oplus G$ *is a sstub-derivation as predicted by Lemma 3.4. Notice that* $d[\mathsf{stub}] \oplus G$ *and* d' *(from Example 3.1) are not compatible, because there is no logical or structural inference rule that can take us from the premise sequent* $A, B \Rightarrow B$ *to the conclusion sequent* $B, G \Rightarrow A \supset B$*.*

Definition 3.3 can be extended in the obvious way to apply to annotated stub-derivations. It is easily verified that Lemma 3.4 holds under substitution of "annotated (s)stub-derivation" for "(s)stub-derivation" in the statement of the lemma.

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

Cut-elimination often involves tracing the "parametric ancestors" (Belnap, 1982) of the cut-formula. The following definition formalizes the tracing of parametric ancestors of a formula occurrence through a derivation. Tracing upwards from the formula occurrence, we will use the symbol $*$ to denote every parametric ancestor except the final (uppermost) parametric ancestor. The symbol $\circ$ is used to denote the final (uppermost) parametric ancestor. That is, an occurrence of the annotated formula A° in a sequent indicates that this occurrence of A is principal in the sequent by the appropriate logical rule or GLR rule.

DEFINITION 3.6 *($f_C[\cdot]$: annotated derivation wrt C). Let τ be a cut-free derivation with endsequent $X \Rightarrow Y$, and C a formula.*

1. *if C is not boxed then let $f_C[\tau] = \tau$.*
2. *if C is boxed ($C \equiv \Box B$) and $\Box B \notin X$ then let $f_{\Box B}[\tau] = \tau$.*
3. *if C is boxed ($C \equiv \Box B$) and $\Box B \in X$. Then τ must be a derivation of the form $\Box B \Rightarrow \Box B$ or $\{\pi\}_1^r / X', \Box B \Rightarrow Y$.*
   Set $f_{\Box B}[\tau]$ as $\Phi_{\Box B}[(\Box B)^* \Rightarrow \Box B)]$ or $\Phi_{\Box B}[\{\pi\}_1^r / X', (\Box B)^* \Rightarrow Y]$ respectively, where $\Phi_{\Box B}$ is defined on the class of cut-free annotated derivations as shown in Table 2.

Observe that the annotation operator $f_C[\cdot]$ is a total function mapping derivations to annotated derivations. Let τ be a derivation with endsequent $X \Rightarrow Y$. If $\Box B \in X$ then the

Form of annotated derivation δ	$\Phi_{\Box B}[\delta]$
$(\Box B)^* \Rightarrow \Box B$	$(\Box B)^\circ \Rightarrow \Box B$
$\dfrac{\{\pi\}_1^r}{\dfrac{G, (\Box B)^{n-1} \Rightarrow H}{G, (\Box B)^{*n} \Rightarrow H}\, LW(\Box B)}$	$\dfrac{\Phi_{\Box B}\left[\dfrac{\{\pi\}_1^r}{G, (\Box B)^{*n-1} \Rightarrow H}\right]}{G, (\Box B)^\circ, (\Box B)^{*n-1} \Rightarrow H}\, LW(\Box B)$
$\dfrac{\{\pi\}_1^r}{\dfrac{G, (\Box B)^{n+1} \Rightarrow H}{G, (\Box B)^{*n} \Rightarrow H}\, LC(\Box B)}$	$\dfrac{\Phi_{\Box B}\left[\dfrac{\{\pi\}_1^r}{G, (\Box B)^{*n+1} \Rightarrow H}\right]}{G, (\Box B)^{*n} \Rightarrow H}\, LC(\Box B)$
$\dfrac{\{\pi\}_1^r}{\dfrac{G', (\Box B)^n \Rightarrow H'}{G, (\Box B)^{*n} \Rightarrow H}\, \rho \neq GLR}$	$\dfrac{\Phi_{\Box B}\left[\dfrac{\{\pi\}_1^r}{G', (\Box B)^{*n} \Rightarrow H'}\right]}{G, (\Box B)^{*n} \Rightarrow H}\, \rho$
$\dfrac{\{\pi\}_1^r}{\dfrac{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}{\Box G, (\Box B)^{*n} \Rightarrow \Box A}\, \text{GLR}}$	$\dfrac{\dfrac{\{\pi\}_1^r}{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\, \text{GLR}$
$\dfrac{\dfrac{\{\pi\}_1^r}{G', (\Box B)^n \Rightarrow H'} \quad \dfrac{\{\pi'\}_1^s}{G'', (\Box B)^n \Rightarrow H''}}{G, (\Box B)^{*n} \Rightarrow H}\, \rho \neq cut$	$\dfrac{\Phi_{\Box B}\left[\dfrac{\{\pi\}_1^r}{G', (\Box B)^{*n} \Rightarrow H'}\right] \quad \Phi_{\Box B}\left[\dfrac{\{\pi'\}_1^s}{G'', (\Box B)^{*n} \Rightarrow H''}\right]}{G, (\Box B)^{*n} \Rightarrow H}\, \rho$

Table 2. *Definition of $\Phi_{\Box B}$. Multisets G and $\Box G$ contain no occurrences of annotated formulae. The function is defined on the class of cut-free annotated derivations. If the antecedent of $ES(\delta)$ does not contain a $(\Box B)^*$ formula occurrence then $\Phi_{\Box B}[\delta] = \delta$. The table shows the other cases.*

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

formula occurrences $(\Box B)^\circ$ and $(\Box B)^*$ in $f_{\Box B}[\tau]$ are each called a *parametric ancestor* of the formula occurrence $\Box B \in X$ in the endsequent. In addition, the formula occurrence $(\Box B)^\circ$ is called the *final parametric ancestor*.

DEFINITION 3.7 *Define $\partial^\circ(B, \tau)$ for a formula B and an annotated derivation τ, as the number of occurrences of the GLR rule in τ whose conclusion contains an occurrence of the annotated formula B° in the antecedent.*

LEMMA 3.8. *Let $d[\mathsf{stub}]$ be an annotated sstub-derivation and G a formula multiset. Then*

(a) $\partial^\circ(B, d[\mathsf{stub}] \oplus G) = \partial^\circ(B, d[\mathsf{stub}])$
(b) Let d' be a derivation such that $d[\mathsf{stub}]$ and d' are compatible. Then $\partial^\circ(B, d[\mathsf{stub}] \leftarrow\!\!\!\shortmid\, d') = \partial^\circ(B, d[\mathsf{stub}]) + \partial^\circ(B, d')$.

Proof.

(a) Because $\partial^\circ(B, \cdot)$ counts the number of instances of the GLR rule with conclusion sequents containing the formula occurrence B°, the result is an immediate consequence of the fact that $\oplus$ does not simply add formulae into the conclusion sequent of an instance of the GLR rule (see Definition 3.3(4)) but inserts an instance of $LW^*(G)$ instead.
(b) By the definition of compatibility, the binding rule for $d[\mathsf{stub}]$ and d' cannot be GLR. Thus if an instance ρ of the GLR rule appears in $d[\mathsf{stub}] \leftarrow\!\!\!\shortmid\, d'$, then ρ must appear in one of $d[\mathsf{stub}]$ or d'. Also, if an instance ρ of the GLR rule appears in either $d[\mathsf{stub}]$ or d', then ρ must appear in $d[\mathsf{stub}] \leftarrow\!\!\!\shortmid\, d'$. The result follows immediately.

□

REMARK 3.9 *Lemma 3.8(a) holds even if G contains decorated formulae.*

DEFINITION 3.10 (width) *Let cut_0 be a topmost cut as shown below:*

$$\dfrac{\dfrac{\{\pi\}_1^r}{X \Rightarrow Y, B}\,\rho \qquad \dfrac{\{\sigma\}_1^s}{B, U \Rightarrow W}}{X, U \Rightarrow Y, W}\,cut_0$$

Then, the width of cut_0 (denoted $width(cut_0)$) is defined as:

$$width(cut_0) = \begin{cases} \partial^\circ(B, f_B[\pi_1]) & \text{if } \rho = GLR \text{ (so } \{\pi\}_1^r = \pi_1) \\ GLR(2, \{\pi\}_1^r / X \Rightarrow Y, B) & \text{otherwise} \end{cases}$$

REMARK 3.11 *Notice that the width has been defined only for a topmost cut as this is sufficient for our purposes. Also, $width(cut_0)$ is independent of the right premise derivation of cut_0.*

EXAMPLE 3.12. *Let us calculate $width(cut_0)$ in the following derivation:*

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box C, C, \Box\Box B, \Box B, \Box B \Rightarrow B}\,GLR \qquad \dfrac{\{\sigma\}_1^s}{\Box D \Rightarrow \Box B}\,LW}{\Box D, \Box\Box B \Rightarrow \Box B}\,L\vee}{\Box C \vee \Box D, \Box\Box B \Rightarrow \Box B}\,LW}{\Box(\Box C \vee \Box D), \Box C \vee \Box D, \Box\Box B \Rightarrow \Box B}\,GLR}{\Box(\Box C \vee \Box D) \Rightarrow \Box\Box B} \qquad \Box\Box B, U \Rightarrow W}{\Box(\Box C \vee \Box D), U \Rightarrow W}\,cut_0$$

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

Writing the left premise derivation of cut_0 *as* $\mu/\Box(\Box C \vee \Box D) \Rightarrow \Box\Box B$, *we get* $width(cut_0) = \partial^{\circ}(\Box\Box B, f_{\Box\Box B}[\mu])$ *where* $f_{\Box\Box B}[\mu]$ *is*

$$\dfrac{\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box C, C, \Box\Box B, \Box B, \Box B \Rightarrow B}}{\Box C, (\Box\Box B)^{\circ} \Rightarrow \Box B}\, GLR \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{\Box D \Rightarrow \Box B}}{\Box D, (\Box\Box B)^{\circ} \Rightarrow \Box B}\, LW}{\Box C \vee \Box D, (\Box\Box B)^{*} \Rightarrow \Box B}\, L\vee}{\Box(\Box C \vee \Box D), \Box C \vee \Box D, (\Box\Box B)^{*} \Rightarrow \Box B}\, LW$$

Because $f_{\Box\Box B}[\mu]$ *contains only one* GLR *rule whose conclusion contains the formula occurrence* $(\Box\Box B)^{\circ}$ *in its antecedent, we have* $width(cut_0) = 1$.

REMARK 3.13 *Let* μ *be the left premise derivation of* cut_0 *in Definition 3.10. Then Valentini (1983, p. 473) defines the width as the cardinality of* $GLR^{(2)}$, *where* $GLR^{(2)}$ *in our notation is the set of all instances* ρ *of* GLR *such that:*

(a) ρ *is a* 2*-ary* GLR *rule in* μ, *and*
(b) B *is the diagonal formula of every* 1*-ary* GLR *rule in* μ *below* ρ, *and*
(c) B *is not introduced by weakening below* ρ.

Definition 3.10 formalizes Valentini's definition, in addition, taking care to account for the interplay of weakening and contraction rules in GLS. Of course, there are no explicit contraction rules in GLS_V *(unlike GLS)—contraction is enforced implicitly as the sequents there are built from sets. The following example illustrates this interplay, and also demonstrates the importance of formalizing the definition, even for* GLS_V.

In the following derivation in GLS, from Definition 3.10, by inspection we get that $width(cut_0) = 1$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box\Box X, \Box X, \Box X, X, \Box\Box C, \Box C, \Box C \Rightarrow C}}{\Box\Box X, \Box X, \Box\Box C \Rightarrow \Box C}\, GLR}{\Box\Box X, \Box X, \Box\Box C, \Box\Box C \Rightarrow \Box C}\, LW(\Box\Box C)}{\Box\Box X, \Box X, \Box\Box C \Rightarrow \Box C}\, LC(\Box\Box C)}{\Box\Box X \Rightarrow \Box\Box C}\, GLR \qquad \Box\Box C, U \Rightarrow W}{\Box\Box X, U \Rightarrow W}\, cut_0$$

In GLS_V, *the sequents* $\Box\Box X, \Box X, \Box\Box C, \Box\Box C \Rightarrow \Box C$, *and* $\Box\Box X, \Box X, \Box\Box C \Rightarrow \Box C$ *are identical because the sequents in that system are built from sets. So the following derivation is legal in* GLS_V:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box\Box X, \Box X, \Box X, X, \Box\Box C, \Box C, \Box C \Rightarrow C}}{\Box\Box X, \Box X, \Box\Box C \Rightarrow \Box C}\, GLR}{\Box\Box X, \Box X, \Box\Box C \Rightarrow \Box C}\, LW(\Box\Box C)}{\Box\Box X \Rightarrow \Box\Box C}\, GLR \qquad \Box\Box C, U \Rightarrow W}{\Box\Box X, U \Rightarrow W}\, cut_0$$

According to Valentini's definition, the width of cut_0 *is* 0 *due to condition (c). If we omit the* $LW(\Box\Box C)$ *rule from the above derivation (observe that the ensuing derivation is still legal in* GLS_V*) then, according to Valentini's definition, the width becomes* 1. *Valentini's arguments in his proof of cut-elimination miss this possibility of a weakening rule whose conclusion sequent is identical to the premise sequent. Indeed, he subsequently argues that whenever the width is* 0 *due to condition (c) then omitting the weakening rule for a*

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

diagonal formula is sufficient to delete the formula from the sequent. The example here shows that this is not the case—indeed, omitting the weakening rule can actually increase the width! So further argument is required in Valentini's proof to handle this case.

This example indicates how the use of explicit contraction rules (through the use of sequents built from multisets) enables us to account accurately for contraction in the sequent. It also demonstrates how the contraction rules require special attention. Thus Moen is certainly justified in asking whether Valentini's arguments can be lifted to multiset-based sequents. However, we will see that Moen's claims about failure of cut-elimination in the new setting are incorrect.

§4. Generalized invertibility of the logical rules. An inference rule in the sequent calculus is called *invertible* if the premise sequents are derivable whenever the conclusion sequent is derivable. The following result is a generalized version of the invertibility result for the logical rules in GLS. We use the term 'generalized' because in cases (iii) and (vi) of the following lemma, the sequents we derive are not exactly the premise sequents of the respective rules. This in turn is due to the fact that we use the 'projective' form of the rules for $L\wedge$ and $R\vee$, rather than the corresponding non-projective form shown below:

$$\frac{A, B, X \Rightarrow Y}{A \wedge B, X \Rightarrow Y} \qquad \frac{X \Rightarrow Y, A, B}{X \Rightarrow Y, A \vee B}$$

LEMMA 4.1. *(general invertibility for logical rules) In the following, derivability is with respect to cut-free derivations in GLS. For all $m \geq 0$,*

(i) *If $(\neg A)^{m+1}, X \Rightarrow Y$ is derivable, then so is $X \Rightarrow Y, A^{m+1}$.*
(ii) *If $X \Rightarrow Y, (\neg A)^{m+1}$ is derivable, then so is $A^{m+1}, X \Rightarrow Y$.*
(iii) *If $(A \wedge B)^{m+1}, X \Rightarrow Y$ is derivable, then so is $A^{m+1}, B^{m+1}, X \Rightarrow Y$.*
(iv) *If $X \Rightarrow Y, (A \wedge B)^{m+1}$ is derivable, then so are $X \Rightarrow Y, A^{m+1}$ and $X \Rightarrow Y, B^{m+1}$.*
(v) *If $(A \vee B)^{m+1}, X \Rightarrow Y$ is derivable, then so are $A^{m+1}, X \Rightarrow Y$ and $B^{m+1}, X \Rightarrow Y$.*
(vi) *If $X \Rightarrow Y, (A \vee B)^{m+1}$ is derivable, then so is $X \Rightarrow Y, A^{m+1}, B^{m+1}$.*
(vii) *If $(A \supset B)^{m+1}, X \Rightarrow Y$ is derivable, then so are $X \Rightarrow Y, A^{m+1}$ and $B^{m+1}, X \Rightarrow Y$.*
(viii) *If $X \Rightarrow Y, (A \supset B)^{m+1}$ is derivable, then so is $A^{m+1}, X \Rightarrow Y, B^{m+1}$.*

Proof. Let us illustrate the proof for (iii) and (vii). The other cases are similar.

Proof of (iii). Suppose that τ is a derivation of $(A \wedge B)^{m+1}, X \Rightarrow Y$. Proof by induction on the height of τ. We must obtain a derivation of $A^{m+1}, B^{m+1}, X \Rightarrow Y$.

First suppose that τ is the initial sequent $A \wedge B \Rightarrow A \wedge B$. Then the following derivation suffices:

$$\dfrac{\dfrac{\dfrac{A \Rightarrow A}{A, B \Rightarrow A}\, LW(B) \qquad \dfrac{B \Rightarrow B}{A, B \Rightarrow B}\, LW(A)}{A, B \Rightarrow A \wedge B}\, R\wedge}$$

Next, consider when $A \wedge B$ is not principal in the last rule ρ in τ (we show when ρ is unary, the binary case is similar). Then τ is of the form:

$$\dfrac{\dfrac{\{\pi\}_1^r}{(A \wedge B)^{m+1}, X' \Rightarrow Y'}\, \rho}{(A \wedge B)^{m+1}, X \Rightarrow Y}$$

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

Notice that it must be the case that $\rho \neq GLR$. By the induction hypothesis we obtain a derivation of $A^{m+1}, B^{m+1}, X' \Rightarrow Y'$. Applying the rule ρ to the sequent $A^{m+1}, B^{m+1}, X' \Rightarrow Y'$ we obtain a derivation of $A^{m+1}, B^{m+1}, X \Rightarrow Y$ as required.

Finally, suppose that $A \wedge B$ is principal in the last rule ρ in τ. If $\rho = L\wedge(A)$ (the case when $\rho = L\wedge(B)$ is similar) then τ has the form:

$$\dfrac{\dfrac{\{\pi\}_1^r}{A, (A \wedge B)^m, X \Rightarrow Y}}{(A \wedge B)^{m+1}, X \Rightarrow Y} L\wedge$$

By the induction hypothesis we obtain a derivation of $A, A^m, B^m, X \Rightarrow Y$. Applying $LW(B)$ we obtain a derivation of $A^{m+1}, B^{m+1}, X \Rightarrow Y$ as required. The remaining possibility to consider is when $\rho = LC(A \wedge B)$. Then τ has the following form:

$$\dfrac{\dfrac{\{\pi\}_1^r}{(A \wedge B)^{m+2}, X \Rightarrow Y}}{(A \wedge B)^{m+1}, X \Rightarrow Y} LC(A \wedge B)$$

By the induction hypothesis we obtain a derivation of $A^{m+2}, B^{m+2}, X \Rightarrow Y$. Applying the rules $LC(A)$ and $LC(B)$ we obtain a derivation of $A^m, B^m, X \Rightarrow Y$ as required.

Proof of (vii). Suppose that τ is a derivation of $(A \supset B)^{m+1}, X \Rightarrow Y$. Proof by induction on the height of τ. We will show how to obtain a derivation of $B^{m+1}, X \Rightarrow Y$. The transformations to obtain $X \Rightarrow Y, A^{m+1}$ are analogous.

If τ is the initial sequent $A \supset B \Rightarrow A \supset B$ then the following derivation suffices:

$$\dfrac{\dfrac{B \Rightarrow B}{A, B \Rightarrow B} LW(A)}{B \Rightarrow A \supset B} R\supset$$

Next, consider when $A \supset B$ is not principal in the last rule ρ in τ (we show when ρ is unary, the binary case is similar). Then τ is of the form:

$$\dfrac{\dfrac{\{\pi\}_1^r}{(A \supset B)^{m+1}, X' \Rightarrow Y'}}{(A \supset B)^{m+1}, X \Rightarrow Y} \rho$$

Notice that it must be the case that $\rho \neq GLR$. By the induction hypothesis we obtain a derivation of $B^{m+1}, X' \Rightarrow Y'$. Applying the rule ρ to the sequent $B^{m+1}, X' \Rightarrow Y'$ we obtain a derivation of $B^{m+1}, X \Rightarrow Y$ as required.

Finally, suppose that $A \supset B$ is principal in the final rule ρ in τ. If it is the case that $\rho = L\supset(A \supset B)$ then τ has the form:

$$\dfrac{\dfrac{\{\pi\}_1^r}{(A \wedge B)^m, X \Rightarrow Y, A} \quad \dfrac{\{\sigma\}_1^s}{B, (A \wedge B)^m, X \Rightarrow Y}}{(A \supset B)^{m+1}, X \Rightarrow Y} L\supset$$

If $m = 0$, then the sequent $B, (A \supset B)^m, X \Rightarrow Y$ is simply $B, X \Rightarrow Y$ so there is nothing more to do. Else, if $m > 0$, by the induction hypothesis applied to the right premise of $L\supset$ we obtain a derivation of $B^{m+1}, X \Rightarrow Y$. Once again, this is the required derivation so there is nothing more to do. The remaining possibility to consider is when $\rho = LC(A \supset B)$. Then τ has the form:

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

$$\dfrac{\dfrac{\{\pi\}_1^r}{(A \supset B)^{m+2}, X \Rightarrow Y}\, LC(A \supset B)}{(A \supset B)^{m+1}, X \Rightarrow Y}$$

By the induction hypothesis we obtain a derivation of $B^{m+2}, X \Rightarrow Y$. Now apply the rule $LC(B)$ to obtain a derivation of $B^{m+1}, X \Rightarrow Y$ as required. □

Note that the above results are not height preserving in general. Embedded inside the induction in the above proof is the notion of tracing the main formula $A \bullet B$ (for $\bullet \in \{\wedge, \vee, \supset\}$) or $\neg A$ upwards from the end-sequent. The presence of the GLR rule does not cause a problem when tracing this formula as it is impossible to encounter, along this path, a GLR rule instance *before* the introduction rule for the principal connective in $A \bullet B$ or $\neg A$. This is because every formula in the conclusion sequent of a GLR rule is necessarily boxed. Notice that proving the result for $m \geq 0$ rather than just for $m = 0$ simplifies matters because we can appeal immediately to the induction hypothesis to handle the case where the main formula is principal due to a contraction rule.

§5. Cut-elimination for *GLS*. The main task for cut-elimination is to show that if $\Box X, X, \Box B \Rightarrow B$ is cut-free derivable in GL, then there is a cut-free derivation of $\Box X, X \Rightarrow B$. This is the content of Lemma 5.3. The cut-elimination theorem follows immediately from this Lemma.

Before proceeding with the technical details let us provide an outline of the proof of Lemma 5.3. Let τ be a cut-free derivation of $\Box X, X, \Box B \Rightarrow B$. Then we define the width $n(\tau)$ as the number of occurrences of the following schema, where no GLR rule occurrences appear between GLR_1 and the endsequent, and $s > 0$:

$$\dfrac{\dfrac{\dfrac{\Box G, G, (\Box B)^s, B^s, \Box C \Rightarrow C}{\Box G, (\Box B)^s \Rightarrow \Box C}\, GLR_1}{\vdots}}{\Box X, X, \Box B \Rightarrow B}$$

If $n(\tau) = 0$ this indicates that the $\Box B$ formula occurrence in the antecedent of the endsequent of τ has either been introduced by $LW(\Box B)$ or can be traced to the initial sequent $\Box B \Rightarrow \Box B$. In the former case, the weakening rule is deleted. In the latter case, the required derivation of $\Box X, X \Rightarrow B$ can be obtained by substituting the derivation $\tau/^{GLR}\Box X \Rightarrow \Box B$ in place of the initial sequent and transforming the sequents below appropriately.

If $n(\tau) = k + 1$, each occurrence of the above schema in τ is transformed so that the GLR_1 rule occurrence is deleted. Observe that the conclusion of the GLR_1 rule can *almost* be obtained by the derivation

$$\dfrac{\Box C \Rightarrow \Box C}{\Box G, (\Box B)^s, \Box C \Rightarrow \Box C}\, LW((\Box G)^s, \Box B)$$

There is now an unwanted $\Box C$ formula occurrence in the antecedent that has to be removed by the appropriate transformations. After $k + 1$ such transformations we obtain the base case of $n(\tau) = 0$.

We are now ready to formalize this argument. We begin with the following decomposition lemma.

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

LEMMA 5.1. *Let τ be a cut-free derivation of the form $\{\pi\}_1^r/^{\rho} X, \Box B \Rightarrow Y$ and suppose that $\partial^\circ(\Box B, f_{\Box B}[\tau]) > 0$.*

(i) If $\rho = GLR$ then $f_{\Box B}[\tau] = \{\pi\}_1^r/^{GLR} X, (\Box B)^\circ \Rightarrow Y$.
(ii) If $\rho \neq GLR$ then we can write the annotated derivation $f_{\Box B}[\tau]$ in the form $d[\mathsf{stub}] \leftarrowtail d'$ such that

$$\partial^\circ(\Box B, d[\mathsf{stub}] \leftarrowtail d') = \partial^\circ(\Box B, d[\mathsf{stub}]) + \partial^\circ(\Box B, d').$$

Furthermore, denote the endsequent of d' as $U \Rightarrow W$. Then for any multiset M, and any derivation d'' with endsequent $U, M \Rightarrow W$, we have that $d[\mathsf{stub}] \oplus M$ and d'' are compatible.

Proof. If $\rho = GLR$ then it follows immediately from Definition 3.6 that $f_{\Box B}[\tau] = \{\pi\}_1^r/^{GLR} X, (\Box B)^\circ \Rightarrow Y$.

Now suppose that $\rho \neq GLR$.

Because $\partial^\circ(\Box B, f_{\Box B}[\tau]) > 0$ the transformed derivation $f_{\Box B}[\tau]$ must have the following form, where

(1) $n \geq 1$, and
(2) $\Box G$ contains no annotated formulae, and
(3) GLR_1 is a 1-ary GLR rule in $f_{\Box B}[\tau]$, and
(4) η may contain branches:

$$\dfrac{\dfrac{\{\pi'\}_1^s}{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\, GLR_1$$
$$\eta$$
$$X, (\Box B)^* \Rightarrow Y$$

We remind the reader that since η is allowed to contain branches, the last rule in η may be either unary or binary. We can identify the annotated derivation $f_{\Box B}[\tau]$ with $d[\mathsf{stub}] \leftarrowtail d'$ where $d[\mathsf{stub}]$ (below left) is an annotated sstub-derivation and d' (below right) is an annotated derivation:

$$\begin{array}{c}\mathsf{stub}\\ \eta\\ X, (\Box B)^* \Rightarrow Y\end{array} \qquad\qquad \dfrac{\dfrac{\{\pi'\}_1^s}{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\, GLR$$

From Lemma 3.8(b) we have

$$\partial^\circ(\Box B, f_{\Box B}[\tau]) = \partial^\circ(\Box B, d[\mathsf{stub}] \leftarrowtail d') = \partial^\circ(\Box B, d[\mathsf{stub}]) + \partial^\circ(\Box B, d').$$

Write the endsequent of d' as $U \Rightarrow W$. Since GLR_1 is a 1-ary GLR rule in $f_{\Box B}[\tau]$ the path (through η) between the leaf stub in $d[\mathsf{stub}]$ and the endsequent $X, (\Box B)^* \Rightarrow Y$ of $d[\mathsf{stub}]$ contains no GLR rule instances. From Definition 3.3 and the compatibility of $d[\mathsf{stub}]$ and d', for any multiset M and any derivation d'' with endsequent $U, M \Rightarrow W$, it follows that $d[\mathsf{stub}] \oplus M$ and d'' are compatible. □

DEFINITION 5.2 (rank of a cut) *For a topmost cut cut_0 the rank $rk(cut_0)$ is the triple (d, n, h) where d is the degree of the cut-formula, n is the width of cut_0, and h is the cut-height of cut_0.*

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

LEMMA 5.3. *Let τ be the following derivation where cut_0 is a topmost cut:*

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\, GLR \qquad \dfrac{\{\sigma\}_1^s}{\Box B, U \Rightarrow W}}{\Box X, U \Rightarrow W}\, cut_0$$

and suppose the following condition holds:

($\star$) *For all derivations δ, every topmost cut in δ with rank $< rk(cut_0)$ is eliminable.*

Then there is a transformed cut-free derivation τ' of $X, \Box X \Rightarrow B$.

Proof. Let μ denote the subderivation $\{\pi\}_1^r / \Box X, X, \Box B \Rightarrow B$ of τ.

Case width($\mathbf{cut_0}$) = 0 : Hence $\partial^\circ(\Box B, f_{\Box B}[\mu]) = 0$. Then the annotated derivation $f_{\Box B}(\mu)$ must have final parametric ancestors of the form $(\Box B)^\circ \Rightarrow \Box B$ or $X' \Rightarrow Y' /^{LW(\Box B)} X', (\Box B)^\circ \Rightarrow Y'$ only.

Let $\Box B^{(*|\circ)}$ stand for an annotated occurrence of $\Box B$ where the annotation is not known. Consider the substitution $(f_{\Box B}[\mu])\{\Box B^{(*|\circ)} := \Box X\}$ obtained by replacing every occurrence

1. of $(\Box B)^*$ with $\Box X$, and
2. of $(\Box B)^\circ \Rightarrow \Box B$ with a derivation of $\Box X \Rightarrow \Box B$ (the cut-free left premise derivation of cut_0), and
3. of $\dfrac{X' \Rightarrow Y'}{X', (\Box B)^\circ \Rightarrow Y'}\, LW(\Box B)$ with $\dfrac{X' \Rightarrow Y'}{X', \Box X \Rightarrow Y'}\, LW^*(\Box X)$

As the endsequent of $f_{\Box B}[\mu]$ was $\Box X, X, (\Box B)^* \Rightarrow B$ we have that the derivation $(f_{\Box B}[\mu])\{\Box B^{(*|\circ)} := \Box X\}$ is a cut-free derivation of $\Box X, X, \Box X \Rightarrow B$. Applying repeated left contraction gives a cut-free derivation of $\Box X, X \Rightarrow B$.

Case width($\mathbf{cut_0}$) > 0 : Hence $\partial^\circ(\Box B, f_{\Box B}[\mu]) > 0$. First suppose that the last rule in μ is GLR. Then μ must be of the form:

$$\dfrac{\dfrac{\dfrac{\{\pi'\}_1^s}{\Box\Box X', \Box X', \Box X', X', \Box\Box B', \Box B', \Box B' \Rightarrow B'}}{\Box\Box X', \Box X', \Box\Box B' \Rightarrow \Box B'}\, GLR}$$

where X' and B' are such that $X = \Box X'$ and $B \equiv \Box B'$. The following is then a derivation of $\Box X, X \Rightarrow B$ where the degree of cut_0 and cut_1 is identical and $width(cut_1) = 0 < width(cut_0)$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Box B' \Rightarrow \Box B'}{\Box B', B', \Box\Box B' \Rightarrow \Box B'}\, LW^*}{\Box B' \Rightarrow \Box\Box B'}\, GLR \qquad \dfrac{\{\pi'\}_1^s}{\Box\Box X', \Box X', \Box X', X', \Box\Box B', \Box B', \Box B' \Rightarrow B'}}{\Box B', \Box\Box X', \Box X', \Box X', X', \Box B', \Box B' \Rightarrow B'}\, cut_1}{\Box\Box X', \Box X', \Box X', X', \Box B' \Rightarrow B'}\, LC^*(\Box B')}{\Box\Box X', \Box X' \Rightarrow \Box B'}\, GLR$$

The required derivation is obtained by using ($\star$) to eliminate cut_1.

If the last rule in μ is not GLR, we can write $f_{\Box B}[\mu]$ as $d[\mathsf{stub}] \leftarrowtail d'$, where $d[\mathsf{stub}]$ and d' (below left and right, respectively) are constructed as in the proof of Lemma 5.1, so $n \geq 1$, $\Box G$ does not contain annotated formulae, and the path in $d[\mathsf{stub}]$ through η from the top-sequent stub to the end-sequent $\Box X, X, (\Box B)^* \Rightarrow B$ contains no occurrence of the GLR rule:

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

$$\dfrac{\dfrac{\mathsf{stub}}{\eta}}{\Box X, X, (\Box B)^* \Rightarrow B} \qquad \dfrac{\dfrac{\{\pi'\}_1^t}{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\, GLR$$

By Lemma 5.1, $\partial^\circ(\Box B, d[\mathsf{stub}] \hookleftarrow d') = \partial^\circ(\Box B, d[\mathsf{stub}]) + \partial^\circ(\Box B, d')$.

Let d'' be the annotated derivation:

$$\dfrac{\Box A \Rightarrow \Box A}{A, \Box A, \Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\, LW^*(A, \Box G, (\Box B)^n)$$

Then $d[\mathsf{stub}] \oplus \{\}A, \Box A$ and d'' are compatible (Lemma 5.1). We observe that $\partial^\circ(\Box B, d') = 1$ and $\partial^\circ(\Box B, d'') = 0$. Recalling that $\lfloor\cdot\rfloor$ is the forgetful map, let Λ_{11} be the derivation:

$$\dfrac{\dfrac{\dfrac{\lfloor d[\mathsf{stub}] \oplus \{A, \Box A\} d''\rfloor}{\Box A, \Box X \Rightarrow \Box B}\, GLR \qquad \dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box A, \Box X, \Box X, X \Rightarrow B}\, cut_1(\Box B)}{\Box A, \Box X, X \Rightarrow B}\, LC^*$$

Let Λ_{12} be the derivation:

$$\dfrac{\dfrac{\dfrac{\lfloor d[\mathsf{stub}] \oplus \{\}A, \Box A \hookleftarrow d''\rfloor}{\Box A, \Box X \Rightarrow \Box B}\, GLR \qquad \dfrac{\{\pi'\}_1^t}{\Box G, G, \Box B, B, \Box A \Rightarrow A}\, LC^*(B^{n-1}, (\Box B)^{n-1})}{\Box A, \Box A, \Box X, \Box G, G, B \Rightarrow A}\, cut_2(\Box B)}{\Box A, \Box X, \Box G, G, B \Rightarrow A}\, LC$$

Consider the derivation Λ_1:

$$\dfrac{\dfrac{\dfrac{\Lambda_{11} \qquad \Lambda_{12}}{\Box A, \Box X, X, \Box A, \Box X, \Box G, G, B \Rightarrow A}\, cut_3(B)}{\Box A, \Box X, X, \Box G, G, \Rightarrow A}\, LC^*}{\Box X, \Box G \Rightarrow \Box A}\, GLR$$

For $i \in \{1, 2\}$, observe that $deg(cut_i) = deg(cut_0)$. Furthermore,

$$\begin{aligned}
width(cut_i) &= \partial^\circ(\Box B, f_{\Box B}(\lfloor d[\mathsf{stub}] \oplus \{\}A, \Box A \hookleftarrow d''\rfloor)) && \text{Def. of width}\\
&= \partial^\circ(\Box B, d[\mathsf{stub}] \oplus \{\}A, \Box A \hookleftarrow d'') && \text{By inspection}\\
&= \partial^\circ(\Box B, d \oplus \{\}A, \Box A[\mathsf{stub}]) + \partial^\circ(\Box B, d'') && \text{Lemma 3.8(b)}\\
&< \partial^\circ(\Box B, d \oplus \{\}A, \Box A[\mathsf{stub}]) + \partial^\circ(\Box B, d')\\
&= \partial^\circ(\Box B, d[\mathsf{stub}]) + \partial^\circ(\Box B, d') && \text{Lemma 3.8(a)}\\
&= \partial^\circ(\Box B, d[\mathsf{stub}] \hookleftarrow d') && \text{Lemma 3.8(b)}\\
&= width(cut_0)
\end{aligned}$$

Because $deg(cut_i) = deg(cut_0)$ and the premises of both cut_1 and cut_2 are cut-free, by appealing twice to $(\star)$ we can in turn eliminate cut_1 and cut_2. In the resulting derivation, since $deg(cut_3) < deg(cut_0)$ we can eliminate cut_3 by $(\star)$. We thus obtain a cut-free derivation Λ_2 of $\Box X, \Box G \Rightarrow \Box A$.

Let Λ_3 be the annotated derivation:

$$f_{\Box B}\left[\dfrac{\Lambda_2}{\Box X, \Box G, (\Box B)^n \Rightarrow \Box A}\, LW^*((\Box B)^n)\right]$$

Clearly $\partial^\circ(\Box B, \Lambda_3) = 0$. Furthermore, by Lemma 5.1, $d[\mathsf{stub}] \oplus \Box X$ and Λ_3 are compatible. Recall that $\lfloor\cdot\rfloor$ is the forgetful map. The endsequent of $\lfloor(d[\mathsf{stub}] \oplus X) \hookleftarrow \Lambda_3\rfloor$ is thus $\Box X, \Box X, X, \Box B \Rightarrow B$. Now consider the derivation Λ_4:

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

$$\dfrac{\dfrac{\dfrac{\dfrac{\lfloor (d[\mathsf{stub}] \oplus \Box X) \leftarrow \Lambda_3 \rfloor}{\Box B, \Box X, X \Rightarrow B}\, LC^*(\Box X)}{\Box X \Rightarrow \Box B}\, GLR \qquad \dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{X, \Box X, \Box X \Rightarrow B}\, cut_4(\Box B)}{X, \Box X \Rightarrow B}\, LC^*(\Box X)$$

By a similar calculation to the above we obtain $width(cut_4) < width(cut_0)$. Because $deg(cut_4) = deg(cut_0)$ and the premises of cut_4 are cut-free, appealing to $(\star)$ we can eliminate cut_4. We thus obtain a cut-free derivation of $X, \Box X \Rightarrow B$. □

Without loss of generality it suffices to consider a derivation concluded by a cut-rule with cut-free premise derivations.

THEOREM 5.4 (Cut-elimination). Let τ be a derivation concluded by an instance cut_0 of the cut-rule with cut-free premise derivations. Then there is a transformed cut-free derivation τ' with identical endsequent.

Proof. Induction on the rank (d, n, h) of cut_0 under the standard lexicographic ordering. We say that the cut-formula is *left principal* if an occurrence of the cut-formula in the succedent of the left premise is a principal formula. The term *right principal* is defined analogously. This follows standard practice.

1 Cut with an initial sequent as premise. This is the base case. The transformations are standard (see Negri & von Plato, 2001; Troelstra & Schwichtenberg, 2000).

2 Cut with neither premise an initial sequent.

(a) Cut-formula is left and right principal.

First suppose that the cut-formula is boxed. There are five possibilities:

(i) the cut-formula is left and right principal by the GLR rule. The derivation must then be in SNF:

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\, GLR \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{\Box B, \Box U, B, U, \Box C \Rightarrow C}}{\Box B, \Box U \Rightarrow \Box C}\, GLR}{\Box X, \Box U \Rightarrow \Box C}\, cut_0$$

The induction hypothesis implies that for any derivation δ, any topmost cut in δ with rank less than $rank(cut_0)$ is eliminable. This is precisely condition $(\star)$ in Lemma 5.3. Hence we can obtain a cut-free derivation of $\Box X, X \Rightarrow B$ from $\{\pi\}_1^r$. Consider the derivation:

$$\dfrac{\dfrac{\dfrac{\Box X, X \Rightarrow B \qquad \dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\, GLR \qquad \dfrac{\{\sigma\}_1^s}{\Box B, \Box U, B, U, \Box C \Rightarrow C}}{\Box X, \Box U, B, U, \Box C \Rightarrow C}\, cut_1}{\Box X, X, \Box X, \Box U, U, \Box C \Rightarrow C}\, cut_2}{\Box X, X, \Box U, U, \Box C \Rightarrow C}\, LC^*}{\Box X, \Box U \Rightarrow \Box C}\, GLR$$

Observe that $rk(cut_1) = (d, n, h - 1)$. By the induction hypothesis we can eliminate cut_1. In the resulting derivation, since $deg(cut_2) < d$, the result follows from another application of the induction hypothesis.

(ii) the cut-formula $\Box B$ is left principal by the GLR rule and right principal by $LC(\Box B)$. Then τ is as below where both premises of cut_0 are cut-free and $m \geq 0$:

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\, GLR \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{(\Box B)^{m+2}, U \Rightarrow W}\,\rho}{\Box B, U \Rightarrow W}\, LC^{m+1}(\Box B)}{\Box X, U \Rightarrow W}\, cut_0$$

In general ρ need not be the GLR rule. However if $\rho \neq GLR$ then either (1) $\rho = LW(\Box B)$ and we delete ρ and the $LC(\Box B)$ rule below it, or (2) $\Box B$ is not principal by ρ.

In case (1) the result is a straightforward application of the induction hypothesis appealing to reduced height. In case (2) the result is obtained by applying ρ *below* cut_0 (as opposed to above the $LC^{m+1}(\Box B)$ rules as it currently stands) and invoking the induction hypothesis. Observe that this is possible even if ρ is a binary rule.

If $\rho = GLR$ it follows that $U \equiv \Box V$ and $W \equiv \Box C$ for some multiset V and formula C, and $s = 1$ and $\sigma_1 \equiv \{\sigma'\}_1^{s'}/(\Box B)^{m+2}, B^{m+2}, \Box V, V, \Box C \Rightarrow C$. Thus τ must be of the form:

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\, GLR \qquad \dfrac{\dfrac{\dfrac{\{\sigma'\}_1^{s'}}{(\Box B)^{m+2}, B^{m+2}, \Box V, V, \Box C \Rightarrow C}\,\rho = GLR}{(\Box B)^{m+2}, \Box V \Rightarrow \Box C}}{\Box B, \Box V \Rightarrow \Box C}\, LC^{m+1}(\Box B)}{\Box X, \Box V \Rightarrow \Box C}\, cut_0$$

A cut-free derivation of $\Box X, X \Rightarrow B$ is obtained as in (i) using Lemma 5.3. Utilizing this derivation, consider the following derivation:

$$\dfrac{\dfrac{\Box X, X \Rightarrow B \qquad \dfrac{\Box X \Rightarrow \Box B \qquad \dfrac{\dfrac{\{\sigma'\}_1^{s'}}{(\Box B)^{m+2}, B^{m+2}, \Box V, V, \Box C \Rightarrow C}}{\Box B, B^{m+2}, \Box V, V, \Box C \Rightarrow C}\, LC^{m+1}(\Box B)}{\Box X, B^{m+2}, \Box V, V, \Box C \Rightarrow C}\, cut_1}{\Box X, B, \Box V, V, \Box C \Rightarrow C}\, LC^{m+1}(B)}{\Box X, X, \Box X, \Box V, V, \Box C \Rightarrow C}\, cut_2$$

$$\dfrac{\Box X, X, \Box V, V, \Box C \Rightarrow C}{\Box X, \Box V \Rightarrow \Box C}\, LC^*$$

$$GLR$$

Now cut_1 has identical degree and width compared to cut_0, and smaller cut-height. Hence, we can eliminate cut_1 using the induction hypothesis. In the resulting derivation $deg(cut_2) < d$ so the result follows from the induction hypothesis.

(iii) the cut-formula $\Box B$ is left principal by $RC(\Box B)$ and right principal by the GLR rule.

Then τ has the following form where both premises of cut_0 are cut-free:

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{X \Rightarrow Y, \Box B, \Box B}}{X \Rightarrow Y, \Box B}\, RC_1 \qquad \dfrac{\sigma_1}{\Box B, \Box U \Rightarrow \Box C}\, GLR}{X, \Box U \Rightarrow Y, \Box C}\, cut_0$$

Because the conclusion of (any) GLR rule has exactly one formula in the succedent, it follows that at least one of the $\Box B$ formula occurrences in the succedent of the premise of RC_1 can be traced upwards in $\{\pi\}_1^r$ to $RW(\Box B)$ rule application(s). In particular, when tracing upwards, it is impossible to encounter a GLR rule application *before* encountering an $RW(\Box B)$ rule application. Deleting these $RW(\Box B)$ rule applications and the RC_1 contraction rule certainly preserves the derivation structure because all the binary rules

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

excluding the cut-rule are additive. This new derivation contains a single instance of cut with identical degree of cut-formula and reduced cut-height compared to cut_0. Furthermore, observe that it must be the case that the width is $\leq n$. The result follows from the induction hypothesis.

(iv) the cut-formula $\Box B$ is left and right principal by $RC(\Box B)$ and $LC(\Box B)$, respectively.

A combination of the strategies in (ii) and (iii) suffice.

(v) the cut-formula $\Box B$ is left or right principal by $RW(\Box B)$ or $LW(\Box B)$, respectively.

Trivial.

When the cut-formula is not boxed and the cut-formula is left and right principal by the respective left and right introduction rule the transformations are standard (see, e.g., Negri & von Plato, 2001; Troelstra & Schwichtenberg, 2000). When the cut-formula is principal in either premise by a contraction rule, we arrive at the well-known case of contractions above cut. For example, in the following derivation, the cut-formula is right principal by $LC(B)$:

$$\dfrac{\dfrac{\{\pi\}_1^r}{X \Rightarrow Y, B}\,\rho \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{B, B, U \Rightarrow W}}{B, U \Rightarrow W}\,LC(B)}{X, U \Rightarrow Y, W}\,cut_0$$

Gentzen (1969) introduced the multicut to handle this case in his proof of cut-elimination for the sequent calculus LK. As noted in the Introduction, we wish to avoid using the multicut rule since it combines the structural rules of contraction and cut. To achieve this we adapt the transformations proposed by von Plato (2001) for the classical calculus $G0c$. Although all the binary rules in $G0c$ are multiplicative, and all the binary rules in GLS are additive, the argument can be lifted here. In particular, the arguments in von Plato (2001) rely on (nonheight preserving) invertibility of the logical rules in $G0c$. It is easily checked that the generalized invertibility results for GLS given in Lemma 4.1 permit a similar argument here.

(b) Cut-formula is left principal only.

(c) Cut-formula neither left nor right principal.

We analyze the last inference rule in the right (resp. left) premise derivation of cut_0. The standard transformations suffice here (see, e.g., Negri & von Plato, 2001; Troelstra & Schwichtenberg, 2000). In particular, observe that for any instance cut_1 of the cut-rule appearing in a transformed derivation, it must be the case that $width(cut_1) \leq n$. □

REMARK 5.5 *In general, it is possible for the width of a lower cut—that is, a nontopmost cut—to increase under the elimination of topmost cuts above it. For example, consider a transformation that turns a topmost cut cut_i (for "initial") into the derivation below containing the cut instance cut_l:*

$$\dfrac{\{\pi\}_1^r \qquad \{\sigma\}_1^s}{G \Rightarrow H}\,cut_l$$

In the case that $\{\pi\}_1^r$ is not cut-free (so cut_l is not a topmost cut) the transformations that ultimately turn cut_l into a topmost cut $cut_{l'}$ could result in a derivation where $width(cut_{l'}) > width(cut_i)$. If this were to occur, then we would be unable to invoke the induction hypothesis to eliminate $cut_{l'}$.

In the proof of Lemma 5.3, cut_4 is the only lower cut that relies on width for elimination. Notice that the above scenario does not occur for cut_4, that is, $width(cut_4)$ does not

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

increase despite the reductions above it. This is because cut_4 in that derivation is 'shielded' by the GLR instance concluding Λ_1.

To see this, observe that derivation Λ_4 in Lemma 5.3 can be written as follows, where η is essentially a cut-free subderivation of the given derivation, and σ is some cut-free derivation:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\sigma}{\Box X, \Box G \Rightarrow \Box A}\, GLR_2}{\Box X, \Box G, (\Box B)^n \Rightarrow \Box A}\, LW^*((\Box B)^n)}{\lfloor \eta \oplus \Box X \rfloor}}{\Box B, \Box X, X \Rightarrow B}\, LC^*(\Box X)}{\Box X \Rightarrow \Box B}\, GLR_1 \qquad \dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{X, \Box X, \Box X \Rightarrow B}\, cut_4(\Box B)}{X, \Box X \Rightarrow B}\, LC^*(\Box X)$$

Crucially, because GLR_2 is a 2-ary GLR rule in the left premise derivation of cut_0, the width of cut_4 is independent of the structure of σ. In other words, if we substituted σ with any other cut-free derivation σ' with identical end-sequent, the width of cut_4 would remain unchanged. This shielding provided by the GLR_2 rule is crucial for the success of the proof.

§6. Moen's Val-II(core) is not Valentini's reduction. We have carefully examined Moen's (2001) slides titled "The proposed algorithms for eliminating cuts in the provability calculus GLS do not terminate."

Although Moen correctly presents Valentini's transformations adapted for the multiset case (he calls this the 'general move VAL-II'), in the remainder of Moen (2001) he never again uses or references this transformation. Instead, Moen uses repeatedly a transformation he variously titles 'Val-II(core)', 'Valentini's (core) algorithm', 'Valentini's algorithm (core)', and 'VAL-II(Core)'. Further evidence that Moen is working with Val-II(core) and not the general move comes from the fact that the move he applies to his 'counterexample' ϵ can easily be checked as being Val-II(core). Thus it is no surprise that the main result of Moen's work 'Theorem 3' states: 'Valentini's algorithm VAL-II (Core) for cut-elimination does not terminate'.

Moen (2001) claims that Val-II(core) is the "...core of Valentini's reduction." Yet Val-II(core) does not appear in Valentini (1983). However it appears in Sambin & Valentini (1982, p. 322) with the comment "this reduction is not sufficient." In other words, Moen has rediscovered what is already known about Val-II(core).

In conclusion, Moen is incorrect in claiming that he has demonstrated that Valentini's algorithm does not terminate—Moen has used the wrong algorithm (i.e., Val-II(core)). In fact, for his concrete derivation ϵ, the width of the cut-formula is 1 so the reduction is straightforward using the transformations given in this paper. Applying the base case transformations, and then the classical transformations, we obtained a cut-free derivation of the end-sequent of ϵ. (Note that there would be little value in presenting the cut-free derivation that was obtained in this manner, since it is the *series of transformations* from ϵ to a cut-free derivation of the identical endsequent that is the focus of this work). Of course, the incorrectness of Moen's claim does not imply the correctness of Valentini's arguments for sequents built from multisets. As we noted in the §1, the whole point is that complications do arise, and these have to be dealt with carefully.

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

§7. Adapting the proof to other logics. An interesting feature of the GLR rule below is that the formula $\Box B$ occurs both in the antecedent of the premise sequent $\Box X, X, \Box B \Rightarrow B$ and in the succedent of the conclusion sequent $\Box X \Rightarrow \Box B$:

$$\frac{\Box X, X, \Box B \Rightarrow B}{\Box X \Rightarrow \Box B}\, GLR$$

Thus, following Sambin & Valentini (1982) we called $\Box B$ the diagonal formula.

More generally we say that a sequent rule ρ has the *diagonal property* if each premise sequent $\mathcal{S}_i$ of ρ has the form $X, f(\Box B_i) \Rightarrow Y$ for multisets X, Y and some formula $f(\Box B_i)$, where (i) $\Box B_i$ occurs positively in $f(\Box B_i)$, and (ii) $\Box B_i$ appears in the succedent of the conclusion of ρ. For any such premise sequent $\mathcal{S}_i$, let $\mathcal{S}_i^-$ denote the sequent $X \Rightarrow Y$ obtained by deleting the $f(\Box B_i)$ formula occurrence. Suppose that $\mathcal{C}$ is a sequent calculus containing a rule ρ with the diagonal property. We say that $\mathcal{C}$ is *ρ-diagonal eliminable* if whenever the premises $\{\mathcal{S}_i\}$ of ρ are cut-free derivable, there is a syntactic transformation deriving the sequents $\{\mathcal{S}_i^-\}$ without using the cut-rule.

Clearly the GLR rule has the diagonal property. Furthermore, Lemma 5.3 shows that GLS is GLR-diagonal eliminable. Recall that this was the main step in the proof of cut-elimination for GLS. In fact, there are several other examples in the literature where the proof of cut-elimination relies on ρ-diagonal eliminability for the modal rule ρ. Define the axioms below:

$$\begin{array}{ll} T: & \Box A \supset A \\ \mathcal{G}rz: & \Box(\Box(A \supset \Box A) \supset A) \supset A \\ \mathcal{G}o: & \Box(\Box(A \supset \Box A) \supset A) \supset \Box A \end{array}$$

and let $L \vdash A$ denote that A is a theorem of the logic L. The logic Grz is obtained from basic modal logic K by the addition of the axiom $\mathcal{G}rz$. The following results are known (see Goré *et al.*, 1997).

$$Grz \vdash T \qquad\qquad Grz \vdash \mathcal{G}o$$

Since $\mathcal{G}rz$ is derivable from $\mathcal{G}o$ and T, an alternative axiomatization for Grz is $KT\mathcal{G}o$. The sequent calculus $GrzS$ (see Avron, 1984) for Grz is obtained by replacing the GLR rule in GLS with the following sequent rules:

$$\frac{B, X \Rightarrow Y}{\Box B, X \Rightarrow Y}\, GRZa \qquad\qquad \frac{\Box X, \Box(B \supset \Box B) \Rightarrow B}{\Box X \Rightarrow \Box B}\, GRZb$$

Notice that $GRZb$ has the diagonal property. Syntactic cut-elimination for Grz is given in Borga & Gentilini (1986). The main lemma in that proof states that if $\Box X, \Box(B \supset \Box B) \Rightarrow B$ is cut-free derivable in $GrzS$, then $\Box X \Rightarrow B$ is cut-free derivable in $GrzS$. In other words, $GrzS$ is $GRZb$-diagonal eliminable. Cut-elimination follows immediately from this result.

Another example is the logic $S4.3Grz$ axiomatized as $KT4.3\mathcal{G}rz$ where:

$$\begin{array}{ll} 4: & \Box A \supset \Box\Box A \\ .3: & \Box(\Box A \supset B) \vee \Box(\Box B \supset A) \end{array}$$

An alternative axiomatization for this logic is $KT4.3\mathcal{G}o$. The sequent calculus $S4.3GrzS$ for $S4.3Grz$ is obtained by replacing the GLR rule in GLS with the following traditional rules:

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

$$\frac{A, X \Rightarrow Y}{\Box A, X \Rightarrow Y}\,\Box \Rightarrow \qquad\qquad \frac{\Box X \Rightarrow A}{\Box X \Rightarrow \Box A}\,\Box_{S4} \Rightarrow$$

and the n-premise sequent rule ($n \geq 1$):

$$\frac{\{\Box X, X, \Box(A_i \supset \Box A_i) \Rightarrow \Box \mathcal{A}_i, A_i\}_{i=1}^{n}}{\Box X \Rightarrow \Box A_1, \ldots, \Box A_n}\; S4.3GrzR$$

where $\mathcal{A}_i = \{A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n\}$. Clearly the $S4.3GrzR$ rule has the diagonal property. Shimura (1991) has presented a proof of cut-elimination for $S4.3GrzS$. The main lemma in that proof shows that if the following derivation is cut-free derivable

$$\frac{\ldots \qquad \dfrac{\{\pi\}_1^r}{\Box X, X, \Box(A \supset \Box A) \Rightarrow \Box Y, A} \qquad \ldots}{\Box X \Rightarrow \Box Y, \Box A}\; S4.3GrzR$$

then there is a syntactic transformation deriving $\Box X, X \Rightarrow \Box Y, A$ without using the cut-rule. This is precisely the statement that $S4.3GrzS$ is $S4.3GrzR$-diagonal eliminable.

As a final example, consider the logic GL_{lin}, axiomatized by the addition of the following axiom to GL:

$$\Box((A \wedge \Box A) \supset \Box B) \vee \Box((B \wedge \Box B) \supset \Box A).$$

Cut-elimination is presented in Valentini (1986) for the sequent calculus $GL_{lin}S$ obtained by replacing the GLR rule in GLS with a sequent rule called LTT. The rule LTT has the diagonal property and, once again, the main part of the proof involves showing that $GL_{lin}S$ is LTT-diagonal eliminable.

In the proof of cut-elimination for each of the logics above, the main difficulty was showing that the sequent calculus was ρ-diagonal eliminable for the sequent rule ρ. Although the technical content of each proof differs, there are some key features and ideas that reappear. It is interesting to investigate what *other* sequent calculi might be addressed using these insights. We observe that a sequent rule with the diagonal property often results when the corresponding logic can be axiomatized using an axiom of a 'similar form' (we use the term informally) to $\mathcal{G}$. Following Demri & Goré (2002), the axioms $\mathcal{G}$ and $\mathcal{G}o$ can be obtained from the schema

$$\Box(\alpha(p) \supset p) \supset \Box p$$

by setting $\alpha(p)$ to $\Box p$ and $\Box(p \supset \Box p)$, respectively. In the remainder of this section, we will consider some other logics that are axiomatized using formulae of a similar form to $\mathcal{G}$.

The logic Go is axiomatized as $K4\mathcal{G}o$. A sequent calculus GoS for Go can be obtained from GLS by replacing the GLR rule with the following sequent rule:

$$\frac{\Box X, X, \Box(B \supset \Box B) \Rightarrow B}{\Box X \Rightarrow \Box B}\; GoR$$

This rule has the diagonal property and so the main difficulty for cut-elimination is showing that GoS is GoR-diagonal eliminable. Our investigations (Goré & Ramanayake, 2012) show that the procedure in this paper can indeed be adapted to achieve this. This leads to a proof of cut-elimination for GoS, although the adaptation is nontrivial and uses a quaternary induction measure (as opposed to the three induction variables used for GLS). Intuitively, instead of working with the 2-ary GLR rules as in GLS, in GoS it becomes necessary to trace the parametric ancestors deeper into the given derivation.

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

Now consider the logics $S4.3.1$ (axiomatized as $KT4.3Dum$) and $S4Dbr$ (axiomatized as $KT4Dbr$, see Goré, 1999), where

$$Dum: \quad \Box(\Box(A \supset \Box A) \supset A) \supset (\Diamond\Box A \supset \Box A)$$
$$Dbr: \quad \Box(\Box(A \supset \Box A) \supset A) \supset (\Box\Diamond\Box A \supset \Box A)$$

Observe the similarity in form of the formulae Dum and Dbr to the $\mathcal{G}$ axiom. Shimura (1992) has shown cut-elimination using a sequent calculus $S4.3.1S$ that requires a cut-free calculus for the logic $S5$ as an oracle. The sequent calculus $S4.3.1S$ is obtained by replacing the $S4.3GrzR$ rule in $S4.3GrzS$ with the rule

$$\frac{S5 \vdash \bigwedge \Box X \supset \Box A_1 \vee \ldots \vee \Box A_n \qquad \{\Box X, X, \Box(A_i \supset \Box A_i) \Rightarrow \Box \mathcal{A}_i, A_i\}_{i=1}^{n}}{\Box X \Rightarrow \Box A_1, \ldots, \Box A_n} \; S4.3.1R$$

where $\mathcal{A}_i = \{\}A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n$. However, using an $S5$ oracle is a rather strong requirement as all the existing cut-free systems for $S5$ are obtained via modification of the traditional sequent calculus, see Sato (1980), Došen (1985), Indrzejczak (1996), and Poggiolesi (2008).

A sequent calculus $S4.3.1S'$ for $S4.3.1$ without using oracles is presented in Goré (1999). This calculus is obtained by replacing $S4.3.1R$ in $S4.3.1S$ with the $2n$-premise sequent rule ($n \geq 1$)

$$\frac{\{\Box X, X \Rightarrow \Box \mathcal{A}_i, \Diamond\Box A_i\}_{i=1}^{n} \qquad \{\Box X, X, \Box(A_i \supset \Box A_i) \Rightarrow \Box \mathcal{A}_i, A_i\}_{i=1}^{n}}{\Box X \Rightarrow \Box A_1, \ldots, \Box A_n} \; S4.3.1R'$$

This rule has the diagonal property. In Goré (1999), a proof of cut-redundancy but not cut-elimination is given for this calculus. It would be interesting to see if $S4.3.1S$ is $S4.3.1R'$-diagonal eliminable. Cut-elimination would follow directly.

To summarize, logics containing an axiom similar in form to the Löb axiom often have a sequent rule ρ with the diagonal property. Very often, proving cut-elimination amounts to showing that the calculus is ρ-diagonal eliminable. Although such proofs are often nontrivial, it is helpful to notice that there are some ideas and features shared with the proof for GLS. Of course, it goes without saying that the success or failure of cut-elimination for any given sequent calculus will depend on the form of the other sequent rules as well.

§8. Conclusion. We have resolved the issue surrounding the use of Valentini's arguments for cut-elimination in a multiset setting for GL. In order to formally define the measure width, we formalized the notion of 'tracing up' a derivation (i.e., identifying the parametric ancestors) via a constructive function. This constructive function can be used to aid the formalization of various other notions in proof theory. Indeed, this function has been used by Dawson & Goré (2010) to machine check Valentini's proof of cut-elimination for GLS_V directly using sequents built from sets, and using multicut to handle the problem of contractions above cut.

In §7 we discussed the possibility of adapting the cut-elimination procedure in this paper, to logics whose axiomatizations contain an axiom of a syntactically similar form to the Löb axiom, and more generally to sequent calculi containing sequent rules with the diagonal property. We are currently working out the details of cut-elimination for such systems in a uniform framework.

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

§9. Acknowledgments. We thank Arnon Avron for his comments on a previous draft of this work.

BIBLIOGRAPHY

Avron, A. (1984). On modal systems having arithmetical interpretations. *Journal of Symbolic Logic*, **49**(3), 935–942.

Avron, A. (1996). The method of hypersequents in the proof theory of propositional non-classical logics. In *Logic: From Foundations to Applications (Staffordshire, 1993)*. Oxford Sci. Publ. New York, NY: Oxford University Press, pp. 1–32.

Belnap, N. D, Jr. (1982). Display logic. *Journal of Philosophical Logic*, **11**(4), 375–417.

Bimbó, K. (2007). $\mathrm{LE}^{\mathrm{t}}_{\rightarrow}$, $\mathrm{LR}^{\circ}_{\overset{\wedge}{\sim}}$, LK, and cutfree proofs. *Journal of Philosophical Logic*, **36**(5), 557–570.

Borga, M. (1983). On some proof theoretical properties of the modal logic GL. *Studia Logica: An International Journal for Symbolic Logic*, **42**(4), 453–459.

Borga, M., & Gentilini, P. (1986). On the proof theory of the modal logic Grz. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, **32**(2), 145–148.

Borisavljević, M., Došen, K., & Petrić, Z. (2000). On permuting cut with contraction. *Mathematical Structures in Computer Science*, **10**(2), 99–136. The Lambek Festschrift: mathematical structures in computer science (Montreal, QC, 1997).

Dawson, J. E., & Goré, R. (2010). Generic methods for formalising sequent calculi applied to provability logic. In Fermüller, C. G., and Voronkov, A., editors. *LPAR (Yogyakarta)*, Vol. 6397 of *Lecture Notes in Computer Science*. Berlin Heidelberg, Germany: Springer, pp. 263–277.

Demri, S., & Goré, R. (2002). Theoremhood-preserving maps characterizing cut elimination for modal provability logics. *Journal of Logic and Computation*, **12**(5), 861–884.

Došen, K. (1985). Sequent-systems for modal logic. *Journal of Symbolic Logic*, **50**(1), 149–168.

Gentzen, G. (1969). The collected papers of Gerhard Gentzen. In Szabo, M. E., editor. *Studies in Logic and the Foundations of Mathematics*. Amsterdam: North-Holland Publishing Co.

Goré, R. (1999). Tableau methods for modal and temporal logics. In *Handbook of Tableau Methods*. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 297–396.

Goré, R., Heinle, W., & Heuerding, A. (1997). Relations between propositional normal modal logics: An overview. *Journal of Logic and Computation*, **7**(5), 649–658.

Goré, R., & Ramanayake, R. (2008). Valentini's cut-elimination for provability logic resolved. In Areces, C., and Goldblatt, R., editors. *Advances in Modal Logic*, Vol. 7. London: College Publications, pp. 91–111.

Goré, R., & Ramanayake, R. (2012). Cut-elimination for weak Grzegorczyk logic *Go*. Accepted subject to minor revisions by Studia Logica.

Guglielmi, A. (2007). A system of interaction and structure. *ACM Transactions on Computational Logic*, **8**(1), Art. 1, 64.

Indrzejczak, A. (1996). Cut-free sequent calculus for S5. *University of Łódź, Department of Logic: Bulletin of the Section of Logic*, **25**(2), 95–102.

Leivant, D. (1981). On the proof theory of the modal logic for arithmetic provability. *Journal of Symbolic Logic*, **46**(3), 531–538.

Mints, G. (2005). Cut elimination for provability logic. In *Collegium Logicum 2005: Cut-Elimination*.

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press

Moen, A. (2001). *The Proposed Algorithms for Eliminating Cuts in the Provability Calculus GLS do not Terminate*. NWPT 2001, Norwegian Computing Center, 2001-12-10. `http://publ.nr.no/3411`.

Negri, S. (2005). Proof analysis in modal logic. *Journal of Philosophical Logic*, **34**(5–6), 507–544.

Negri, S., & von Plato, J. (2001). *Structural Proof Theory*. Cambridge, UK: Cambridge University Press. Appendix C by Aarne Ranta.

Poggiolesi, F. (2008). A cut-free simple sequent calculus for modal logic S5. *The Review of Symbolic Logic*, **1**(1), 3–15.

Poggiolesi, F. (2009). A purely syntactic and cut-free sequent calculus for the modal logic of provability. *Review of Symbolic Logic*, **2**(4), 593–611.

Pottinger, G. (1983). Uniform, cut-free formulations of T, S4 and S5 (abstract). *Journal of Symbolic Logic*, **48**, 900–901.

Sambin, G., & Valentini, S. (1982). The modal logic of provability. The sequential approach. *Journal of Philosophical Logic*, **11**(3), 311–342.

Sasaki, K. (2001). Löb's axiom and cut-elimination theorem. *Journal of the Nanzan Academic Society of Mathematical Sciences and Information Engineering*, **1**, 91–98.

Sato, M. (1980). A cut-free Gentzen-type system for the modal logic S5. *Journal of Symbolic Logic*, **45**(1), 67–84.

Shimura, T. (1991). Cut-free systems for the modal logic S4.3 and S4.3Grz. *Reports on Mathematical Logic*, **25**, 57–73.

Shimura, T. (1992). Cut-free systems for some modal logics containing S4. *Reports on Mathematical Logic*, **26**, 39–65.

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

Švejdar, V. (1999). On provability logic. *Nordic Journal of Philosophical Logic*, **4**(2), 95–116.

Takeuti, G. (1987). *Proof Theory* (second edition), Vol. 81 of *Studies in Logic and the Foundations of Mathematics*. Amsterdam: North-Holland Publishing Co.

Troelstra, A. S., & Schwichtenberg, H. (2000). *Basic Proof Theory* (second edition), Vol. 43 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge, UK: Cambridge University Press.

Valentini, S. (1983). The modal logic of provability: Cut-elimination. *Journal of Philosophical Logic*, **12**(4), 471–476.

Valentini, S. (1986). A syntactic proof of cut-elimination for GL_{lin}. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, **32**(2), 137–144.

von Plato, J. (2001). A proof of Gentzen's Hauptsatz without multicut. *Archive for Mathematical Logic*, **40**(1), 9–18.

LOGIC AND COMPUTATION GROUP
RESEARCH SCHOOL OF COMPUTER SCIENCE
THE AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, ACT 0200, AUSTRALIA
E-mail: rajeev.gore@anu.edu.au

CNRS LIX, ÉCOLE POLYTECHNIQUE
91128 PALAISEAU, FRANCE
E-mail: revantha@lix.polytechnique.fr

https://doi.org/10.1017/S1755020311000323 Published online by Cambridge University Press